Adverse Selection in Prediction Markets: Evidence from Kalshi*

Robert Bartlett† Maureen O'Hara‡

This version: July 2, 2026 (first version: April 16, 2026)

Abstract

Using 41.6 million trades, we measure adverse selection in prediction markets. Adapting Kyle's lambda and the Glosten-Harris decomposition, we show single-name markets exhibit greater informed price impact than broad-based markets. Despite this, effective spreads are only modestly wider, and market makers earn twice as much per contract. A frequency-magnitude decomposition resolves this puzzle: traders systematically overbet YES in markets that predominantly settle NO, generating a behavioral surplus that cross-subsidizes adverse selection. Adapting the VPIN toxicity metric, we find that one-sided order flow predicts maker losses in single-name markets. Our research suggests a new microstructure equilibrium for binary settlement markets.

JEL Codes: G10, G14, G23
Keywords: prediction markets, adverse selection, informed trading, market microstructure, event markets

*We thank Ilya Beylin, David Easley, Joe Grundfest, and Albert Menkveld for helpful comments.
†Stanford Law School; Stanford Graduate School of Business. Email: rbartlett@law.stanford.edu.
‡Cornell University, Johnson College of Business. Email: mo19@cornell.edu.

1 Introduction

On February 9, 2026, Kalshi (the first CFTC-regulated prediction market exchange) settled the event "Will Jeff Bezos attend the Super Bowl?" at NO. Over the preceding week, 888,006 contracts had changed hands on this single question. The *Wall Street Journal* later identified a source of informed trading: Evan Whitesell, Bezos's stepson and a member of Sigma Alpha Epsilon at the University of Miami mentioned his family's Super Bowl plans at a fraternity event, and fraternity brothers traded on the information.[1] The YES price fell from roughly 70 cents to 30 cents as fraternity networks bought NO. This is, by any definition, informed trading: someone with private knowledge of the event outcome was on the other side of the book from the liquidity providers.

This episode raises a question that is central to the economics of market making in prediction markets: can the adverse selection framework developed for equity markets (Glosten and Milgrom [1985], Kyle [1985], and the extensive empirical literature that followed [Glosten and Harris, 1988, Hasbrouck, 1991, Huang and Stoll, 1996]) be applied to these new markets? And if informed trading is present, as the Bezos event market suggests, why do prediction market makers appear to do so well? The astonishing growth of Kalshi, particularly over the last year, underscores the importance of understanding these vibrant new markets (see Figure 1).

The Bezos incident is not an isolated curiosity. Kalshi has sanctioned traders for exploiting nonpublic information on at least two occasions (a political candidate who bet on his own race and a media employee who traded on embargoed content), and the exchange's Super Bowl markets alone generated over $1 billion in volume amid widespread discussion of insider trading risk.[2] But Kalshi also lists hundreds of thousands of markets on macroeconomic

[1] "The Wildest Frat Party on Campus: Prediction Markets," *Wall Street Journal*, March 5, 2026.

[2] "Kalshi Found Some Insider Traders," Matt Levine, *Bloomberg Opinion*, February 25, 2026. See also "Predicting the Big Game," Matt Levine, *Bloomberg Opinion*, February 9, 2026; "Prediction markets allow trading on Super Bowl commercials, prompting insider trading questions," *CNBC*, January 30, 2026. For a systematic analysis of informed trading across prediction market platforms, see Mitts and Ofir [2026], who document cases spanning the 2026 Iran strike, corporate announcements, and celebrity events. The CFTC issued its first prediction market insider trading advisory in February 2026 (Press Release No. 9185-26).

indicators, market indices, and asset prices where the risk of informed trading would appear to be far less acute. No individual possesses a decisive private signal about where the S&P 500 will close tomorrow or what the next CPI print will be. The juxtaposition of event markets with concentrated private information and markets with dispersed public information, all trading on the same exchange through the same order book, creates a natural setting for studying how varying levels of adverse selection affect market making in prediction markets.

We exploit this variation using the trading record from Kalshi, drawing on 41.6 million trades across 478,167 markets that we classify as either *single-name* (referencing a specific company or individual) or *broad-based* (referencing macroeconomic aggregates, market indices, or asset prices).[3] Kalshi's data offer two features that are, to our knowledge, unavailable in any equity market dataset. First, every trade record discloses the taker side (whether the party crossing the spread bought YES or NO), eliminating the need for trade-classification algorithms. Second, every market settles at \$1 (YES) or \$0 (NO), so the realized profit or loss for the maker on every trade can be computed exactly. And relative to other prediction markets such as Polymarket, whose public records contain only executed trades, Kalshi also publishes quoted bid and ask prices, so the spread that canonically compensates adverse selection can be measured directly rather than inferred.

The parallel to single-stock versus index-ETF trading in equity markets [Madhavan, 2002] is direct: concentrated information environments should produce more adverse selection. It should be stressed that we are focusing here on informed trading, which may arise from either insider trading or simply reflect a trader's own research leading to a more informed view on the issue at hand. Both elements of informed trading may be more likely in a concentrated information environment. Single-name and broad-based markets trade on the

[3]Kalshi lists over 7 million markets spanning politics, sports, weather, entertainment, and other categories, though 87% are programmatically generated parlay markets that bundle combinations of the roughly 1 million standalone markets into multi-leg bets. We restrict the main analysis to single-name and broad-based standalone event markets where the cross-sectional variation in information concentration should be especially acute (see Internet Appendix A for details on how we classify Kalshi markets). Internet Appendix B tests this conjecture by estimating our adverse selection metrics across all non-parlay Kalshi market categories. Single-name markets exhibit the highest informed price impact of any category and broad-based among the lowest, with politics, sports, entertainment, weather, and other markets falling in between.

same exchange, through the same order book, with the same population of takers and makers. Only the information structure differs, raising the prospect that single-name markets exhibit more adverse selection.

Adapting Kyle's λ to the prediction market setting using a log-odds transformation [Tsang and Yang, 2026], we find that they do. Single-name markets exhibit significantly higher informed price impact ($t = 10.88$, $p < 0.001$), and a Glosten-Harris decomposition [Glosten and Harris, 1988] confirms that this impact is almost entirely permanent (information, not inventory), with the transitory component economically negligible. These results hold with and without controls and across activity levels.

Given these findings, canonical adverse selection models predict wider spreads, and we do find wider effective spreads on single-name markets when measured at the freshest available quotes. But the canonical prediction does not stop there. It holds that the wider spread also *compensates* the maker, the markup earned from uninformed traders offsetting the losses to the informed, the balance a Glosten and Milgrom [1985] dealer strikes by completing the round trip and ending flat. That round trip is available to a prediction-market maker too, but she typically does not complete it; she holds her position to settlement, an empirical regularity we document. The spread cannot protect a position held that way: an informed trade then costs the maker the full distance to the \$0 or \$1 outcome, not the few cents of adverse drift typical of equities, and the thin spread a competitive book sustains covers only a fraction of it. Carried to its conclusion, the standard theory predicts that makers should withdraw from precisely the single-name markets where adverse selection is most severe. They do not. The question is therefore not whether spreads widen but what sustains maker profitability where the spread cannot.

The answer lies in how makers actually make money, and this is where the prediction market setting is genuinely illuminating.[4] Because we observe both the taker side and the

[4]Kalshi's maker population is heterogeneous by design. The exchange runs two tiers of liquidity provision: a gated Market Maker Program, under which designated firms accept two-sided quoting obligations in exchange for reduced fees and adjusted position limits (the *Wall Street Journal* reports Susquehanna International Group as the first such firm), and an open Liquidity Incentive Program that rewards ordinary

settlement outcome for every trade, we can decompose maker profits into two components: a *frequency edge* (winning on more contracts than losing) and a *magnitude edge* (winning more per contract than losing). On single-name markets, the frequency edge is the maker's entire source of profit and then some: makers win on 63.4% of contracts, but when they lose, they lose more per contract (23.13 cents) than they earn when they win (16.34 cents). The frequency edge (+5.31 cents) must overcome a magnitude *deficit* (−3.40 cents) to produce a net profit of 1.91 cents per contract.

Broad-based markets, where takers are better calibrated, tell a similar story at smaller scale. Makers earn 0.82 cents per contract, with the frequency edge (+2.01 cents) again subsidizing a magnitude deficit (−1.19 cents). Makers earn over twice as much per contract on single-name markets, where adverse selection is most severe.

What makes this equilibrium work is a pervasive miscalibration between the prices single-name takers pay and the rates at which events actually settle. Across the entire probability spectrum, single-name YES prices exceed realized YES-settlement rates, reflecting a systematic YES optimism that differs from the classical favorite-longshot bias in prior studies of betting markets [Wolfers and Zitzewitz, 2004] in that it is not confined to extreme probabilities but is present even at mid-range prices. Becker [2025] has documented a related "Optimism Tax" on Kalshi from the taker's perspective, and Bürgi, Deng, and Whelan [2026] document the favorite-longshot bias on the same exchange; our contribution is to show that this behavioral surplus is concentrated in single-name markets, precisely where

members for posting resting orders that improve the book. Consistent with this mixture, the exchange notes that only about "5% of bid matches on the platform are from the big institutional market makers you think of," the rest coming from "over 2,000 smaller market makers, some of which are just individuals" [Butcher, 2025]. Over our sample period the large equity market-making firms (Citadel Securities, IMC, and Hudson River Trading) stayed away, deterred in part by the absence of any instrument with which to hedge a held position, while the institutional makers that did enter, Susquehanna foremost, came from sports-betting rather than equity-market-making backgrounds ("Event Bets Pose a Problem for Wall Street Firms Looking to Trade," *Bloomberg*, April 12, 2026). Professional firms and small retail makers thus coexist in the same book, and the mechanism we document does not turn on which of them provides liquidity, since the surplus accrues to whomever holds the resting orders when a market settles. Nor can rebates account for the maker profits we measure: the Liquidity Incentive Program excludes designated market makers and pays only a small fixed daily pool on flagged markets, which is negligible against the behavioral surplus we document and which is accumulated across millions of single-name trades.

adverse selection is highest, and that it is large enough to cover the losses that informed traders impose on makers. Their analyses document taker biases but contain no informed trading, and so cannot address the role it plays in sustaining liquidity in the very markets where adverse selection is most severe.

Because the market is dynamic, a challenge for the market maker is to be able to recognize that informed trading can be time-varying, leading to an increase in trade toxicity. To investigate whether informed trading can be detected in real time, we adapt the VPIN flow toxicity metric [Easley, López de Prado, and O'Hara, 2012] to the prediction market setting. Standard equity VPIN requires several modifications for binary contract markets, including a restriction to interior prices (30–70 cents) to address a settlement-outcome endogeneity that has no equity-market analog: as a contract's outcome becomes clear and its price converges to 0 or 100, flow becomes mechanically one-sided, inflating VPIN without reflecting informed trading. At interior prices, where this confound is absent, trailing VPIN predicts maker losses on single-name markets (OLS $t = -2.42$; logit $z = 3.58$) but not for broad-based (OLS $t = 0.56$; logit $z = 0.10$). The divergence reflects a fundamental difference in what one-sided flow means across the two contract types. On single-name markets, the direction of flow predicts the settlement outcome: among the highest-VPIN buckets, those with YES-directed flow settle YES significantly more often than those with NO-directed flow ($t = 12.44$). On broad-based markets, flow direction and settlement outcome are unrelated ($t = -0.62$), because one-sided flow on a Bitcoin or S&P 500 contract reflects the recent trajectory of a continuously moving price that can reverse before settlement. Consistent with the maker profitability decomposition, the VPIN effect on single-name markets is concentrated on NO-settling markets ($t = -1.99$), where informed NO flow erodes the optimism-tax profits that fund the maker's operations.

Taken together, these findings describe an equilibrium with no equivalent in equity markets. The maker holds the positions she takes to settlement and profits, in the aggregate, when the price she was paid exceeds the realized outcome, and what makes holding prof-

itable is the behavioral surplus, the systematic overpayment of YES-optimistic takers. What is remarkable is that this surplus scales with the adverse selection it must offset: it is largest in exactly the single-name markets where informed trading is most severe, so spreads need not widen to compensate for that risk, keeping trading costs low and volume plentiful even there. The market is in this sense self-reinforcing, since the concentrated-information markets that draw informed traders are the same ones that draw the optimistic flow that pays to absorb them. This is what the single-name and broad-based comparison reveals, beyond both the longshot-bias literature and the textbook prediction that adverse selection rises with information concentration. The equilibrium has, moreover, proven durable as Kalshi has scaled.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature and Section 3 describes the Kalshi exchange, discusses a microstructure framework for prediction market analysis, and sets out our data. Sections 4 presents our adverse selection findings, while Section 5 analyzes maker profitability across single-name and broad-based markets. Section 6 adapts VPIN to the prediction market setting. Section 7 discusses implications for microstructure modelling and for the regulation of informed trading, and Section 8 concludes. Internet Appendix A details our market classification methodology, and Internet Appendix B extends the adverse selection and calibration analyses to all non-parlay Kalshi market categories, confirming that single-name and broad-based markets sit at opposite ends of the informed trading spectrum.

2 Related Literature

The idea that some traders know more than others, and that this asymmetry shapes the cost of trading, is foundational to market microstructure. Copeland and Galai [1983], Glosten and Milgrom [1985], and Kyle [1985] formalize the problem: a market maker who cannot distinguish informed from uninformed order flow must set a spread wide enough to break even

against the informed, and the resulting adverse selection cost is borne by all traders. The empirical literature that followed developed increasingly sophisticated tools for detecting information in trade data. Glosten and Harris [1988] decompose spreads into permanent and transitory components; Hasbrouck [1991] measures the information content of trades through their permanent price impact; Easley, Kiefer, O'Hara, and Paperman [1996] develop the PIN model to estimate the probability that a given trade originates from an informed counterparty. For comprehensive treatments, see O'Hara [1995] and Hasbrouck [2004].

A central prediction of this literature is that adverse selection should be more severe when private information has greater precision. In equity markets, the evidence is consistent with this prediction. For instance, single stocks have wider spreads and higher price impact than index ETFs [Madhavan, 2002], and firms with greater insider ownership or less analyst coverage exhibit elevated adverse selection costs. Our paper tests the analogous prediction in a setting where the cross-sectional variation in information concentration is, if anything, even starker than in equities. Broad-based markets can be affected by a variety of factors; single events by far fewer. Consequently, we hypothesize that adverse selection should be of greater importance in single-event markets.

Economists have long been interested in prediction markets, with early work focusing on gambling on horse races (see, for example, Dowie 1976 and Ali 1977). Using British racing data, Dowie [1976] established empirically the favorite-longshot bias in which bettors tend to overvalue longshots and undervalue favorites, a finding that has carried over to most prediction market settings [Wolfers and Zitzewitz, 2004]. Shin [1991, 1992] analyzed the pricing problem of a monopolist bookmaker who is himself uninformed but faces bettors with inside information, working in the setting of Arrow-Debreu securities. He showed that such a bookmaker optimally shades the odds to protect against adverse selection, and that the favorite-longshot bias arises as the signature of this protection rather than from any pre-existing bias among bettors, since absent insiders the equilibrium odds coincide with the true probabilities. When the incidence of insider trading becomes large enough, the bookmaker

withdraws and the market ceases to function [Shin, 1991]. Levitt [2004] investigates sports betting on NFL games, where a bookmaker sets a point spread and bettors choose which side of the bet to take. He too finds that a skewed pricing strategy can be optimal, but only if the bookmaker has market power, and in his setting the bookmaker relies on a set of specialized odds-makers whose forecasts are at least as good or better than those of any bettor. A defining feature of his setting is the absence of adverse selection, since he finds no evidence that bettors systematically beat the bookmaker, so that the bookmaker's profits come from exploiting a behavioral bias rather than from defending against better-informed traders.

In more standard microstructure models, and in the Kalshi setting we investigate here, the market maker is not presumed to have such expertise but rather faces sure losses when trading with agents who possess superior information. Market makers are likewise not presumed to have market power, so a maker's pricing cannot emulate that of the bookmaker in either Levitt's or Shin's account. Given these differences from gambling markets, Levitt surmises that in a financial market, where the prevalence of inside information prevents any one participant from forecasting better than the market, the best a maker can do is to collect the bid-ask spread. We show that this conjecture does not hold in the Kalshi markets we analyze. An analogous pricing miscalibration arises even though no participant occupies the bookmaker's position, and because it is systematic it is also predictable. A maker need not forecast any single outcome better than the market, which Levitt's premise denies her, to earn more than the spread; it is enough that the takers' overpayment is predictable in the aggregate. She collects the bid-ask spread where two-sided flow permits and, on the positions she carries to settlement, harvests that predictable behavioral subsidy. Makers thus remain viable in a competitive limit order book, none of them able to shade the odds, not by pricing like a bookmaker but by holding against a bias they can anticipate. Our analysis sheds new light on how a prediction market sustains liquidity without the institutional assumptions on which the prior theory rests.

Prediction markets have also attracted sustained attention as information aggregation mechanisms. Wolfers and Zitzewitz [2006] survey the evidence that market prices in betting markets outperform polls and expert surveys, and Arrow [2008] advocates broader adoption. Swanson, Wang, and Wu [2025] show that Kalshi macro markets rival traditional forecasts for economic variables. Bürgi, Deng, and Whelan [2026] using Kalshi data identify the longshot bias and heterogeneous beliefs in affecting profit and loss outcomes. Becker [2025] reaches similar conclusions showing that these effects are strongest for contracts on sports and entertainment. Both papers focus on behavioral explanations for wealth transfers. Palumbo [2026] argues that risk allocation is more akin to underwriting than to the intermediation generally found in trading markets. None of these papers measure adverse selection or identify which contracts are most susceptible to informed trading.

A nascent literature has begun to apply microstructure tools to prediction market data, and the early results are suggestive. Tsang and Yang [2026] adapt Kyle's λ to Polymarket election markets using a log-odds transformation, showing that assassination attempts produce permanent price revisions (information) while debate outcomes produce transitory effects (noise). Hauser and Rittler [2024] apply the Kyle (1985) model to a large prediction market dataset, identifying traders with positive permanent price impact who earn informed-trader profits. Ng, Peng, Tao, and Zhou [2026] find that net order imbalance from large trades predicts returns across multiple platforms. Separately, Mitts and Ofir [2026] document specific cases of apparent informed trading on Polymarket and Kalshi, screening over 210,000 suspicious wallet-market pairs across geopolitical and corporate events and estimating approximately $143 million in aggregate anomalous profit.

Our paper explores the extent to which the standard tools of equity market microstructure can be used to assess adverse selection in prediction markets. Facilitating our research are two features of Kalshi's data. First, the taker side field on every trade eliminates the classification error that plagues equity-market adverse selection estimates, allowing clean measurement of

signed order flow without the Lee and Ready [1991] algorithm.[5] Second, because we observe both the taker side and the settlement outcome, we can compute the realized profit or loss on every trade for both maker and taker, decompose the sources of maker profitability, and explain why spreads need not widen even when adverse selection is severe.

3 Institutional Setting and Data

3.1 The Kalshi Exchange

Kalshi is a CFTC-regulated designated contract market (DCM) that lists binary event contracts. Each binary contract pays \$1 if a specified event occurs and \$0 otherwise; prices are quoted in cents (1–99¢), which are directly interpretable as market-implied probabilities under risk neutrality.

What distinguishes Kalshi from its predecessors is not the contract design (PredictIt and the Iowa Electronic Markets offered similar binary payoffs) but the trade data. Earlier platforms operated continuous limit order books but did not record which side of each trade was the aggressor. Kalshi, by contrast, operates a central limit order book that explicitly tags each trade with the identity of the aggressor. Every trade record discloses the taker side (whether the party who crossed the spread bought YES or NO contracts), and the non-aggressing counterparty is the maker whose resting limit order was executed. This single field eliminates the need for trade-classification algorithms.

Both sides of every Kalshi trade are fully collateralized at execution: the YES buyer deposits the trade price, the NO buyer deposits the complement (\$1 minus the trade price), and the exchange holds these funds in segregated accounts at regulated financial institutions. At settlement, Kalshi distributes \$1 per contract to the winning side and zero to the losing side. There is no counterparty credit risk: the winner's payout is guaranteed by collateral

[5] See also Easley, López de Prado, and O'Hara [2016] for discussion of alternative methods of trade classification.

already posted, not by the loser's ability to pay.[6]

3.2 Market Making in a Prediction Market

3.2.1 A Microstructure Framework

Like any financial market, a prediction market faces the challenges of liquidity and price discovery, and how this is accomplished depends on the microstructure of the market. The central limit order book structure of Kalshi means that liquidity will be provided by those market participants willing to post orders in the book. As in any limit order market, these orders can come from specialized intermediaries or they can come from traders who wish to trade but who do not want to cross the spread. For the purposes of our discussion here we will refer to both groups of liquidity providers as market makers.

Microstructure models generally view traders in financial markets as being either informed traders or liquidity traders, and both categories require recalibration in this setting. In standard microstructure models liquidity traders trade for exogenous reasons such as portfolio rebalancing, and so their trades are not information-related. In a prediction market no one is trading to rebalance a portfolio, so this motivation for liquidity trading is not useful. Liquidity traders here are instead those who have a belief about the outcome of a particular contract but no specific information about the final outcome (i.e., they are not Jeff Bezos's stepson). These liquidity traders are akin to the "quasi-informed" traders proposed by Bagehot [1971], who think they know something about the future outcome but really do not, and who, Bagehot argued, provide the liquidity against which informed traders trans-

[6]This contrasts with traditional over-the-counter derivatives, where counterparty risk is managed through margin calls, netting agreements, and central clearing. On Kalshi, the maximum loss is bounded by the initial collateral, and the exchange cannot be called upon to cover defaults. Full collateralization is a condition of Kalshi's regulatory approval. Kalshi Klear LLC, a CFTC-registered derivatives clearing organization (DCO), is required to clear only fully collateralized positions, as defined in Commission Regulation 39.2: the clearing organization must hold, at all times, funds sufficient to cover the maximum possible loss on every position. Kalshi's clearing system enforces this through a pre-trade "cap check" that validates sufficient collateral at the order stage, before execution, rejecting any order for which the member's account lacks sufficient funds. See Kalshi Klear LLC Order of Registration as a DCO (August 28, 2024); KalshiEX Immediate Account Funding Filing, CFTC Rule 40.6(a) (December 2021).

act. On Kalshi these quasi-informed traders are not symmetric; as Becker [2025] shows, they are systematically YES-optimistic, overpaying for YES contracts, which underperform their NO counterparts across the price range. Informed traders, by contrast, are characterized as receiving a signal about the actual outcome, and market makers lose when they trade with them, which is the adverse selection this paper investigates. The informed can include those with inside information and those whose knowledge gives them a betting advantage over uninformed traders, and the line between quasi-informed and informed traders can be hard to delineate.

In practice, sorting traders into specific categories is not fruitful, but sorting them by trading behavior is. Informed traders will all be on one side of the market and will continue to trade until the price reaches its "true" value, whereas uninformed traders can be on both sides of the contract and may or may not continue to trade. Uninformed trades can be autocorrelated due to algorithms splitting large trades into multiple pieces, or due to correlated noise trading inducing sequences of one-sided trades. These uninformed trade sequences make it harder for the market maker to learn from order flow, and in the worst case could vitiate its information content. The risk would appear especially acute in prediction markets given the YES-optimism we document, which raises the prospect of correlated YES-based noise trading moving prices without actually carrying any information about settlement outcomes. At the same time, a distinctive feature of the prediction markets we study is that we can examine whether observed one-sided flow was on the correct side of the eventual outcome, which allows us to decompose one-sided flow into that which correctly predicts outcomes and that which does not. As we show in Internet Appendix B, the impact on prices associated with one-sided flow on the wrong side of settlement is overwhelmed by the impact on prices associated with one-sided flow on the correct side of settlement. Trade imbalance is thus a noisy but effective indicator of information-based trading.

Market makers set bid and ask prices to balance the losses to informed traders against the gains from liquidity traders, so knowing the "toxicity" of the order flow is important to their

decision about where and when to place limit orders. Easley, Kiefer, O'Hara, and Paperman [1996] developed the PIN model to estimate this probability of informed trade, with the VPIN measure of Easley, López de Prado, and O'Hara [2012] providing dynamic estimates.[7] As we discuss later in the paper, toxicity estimates play a large role in algorithmic trading, adjusting or even curtailing a market maker's liquidity provision. If toxicity estimates are informative, market makers can better predict the ending value of each contract. A question we ask in this paper is whether market makers can estimate the toxicity of order flow in real time in the Kalshi prediction market.

3.2.2 Compensating the Liquidity Provider

Whatever its source, adverse selection risk must be compensated or makers will withdraw. The canonical channel is the bid-ask spread. An equity maker buys at the bid and sells at the ask, ending flat, and the spread is set wide enough that the markup on uninformed trades covers the losses to the informed [Glosten and Milgrom, 1985]. Inventory models broaden the rationale for the spread to include compensation for bearing price risk while the maker waits for the other side of the trade to arrive [Grossman and Miller, 1988, Ho and Stoll, 1981], but the spread is still the vehicle, and the maker still earns by buying low and selling high, only with a delay and some risk.

Yet compensation for the maker need not arrive this way. A liquidity provider can instead be paid through the price of the position itself rather than a round-trip markup. For instance, when a position cannot be hedged, end-user demand pushes its price away from fair value, and the provider who supplies the exposure earns that premium for bearing the unhedgeable risk [Gârleanu, Pedersen, and Poteshman, 2009]. Carried to its limit, she does not round-trip at all. She takes a position, holds it to resolution, and profits if the price she was paid

[7] Like Kyle's λ, VPIN is non-directional by construction, measuring the intensity of one-sided flow rather than its sign, so it too cannot by itself separate informed flow from the directional noise that systematic YES-optimism can generate. We apply the same settlement-conditioning used in the decomposition above: in Section 6 we show that the direction of one-sided flow predicts the settlement outcome on single-name markets but not on broad-based ones, separating toxicity that reflects private information from one-sided flow that merely tracks a moving underlying.

exceeds the realized value, the way an insurer earns a premium for bearing risk to expiry. Her expected profit then comes not from the spread but from the settlement outcome.

Several features of a prediction market push its makers toward this second model. The first is the composition of order flow. Informed traders share a view of where the contract will settle, so their orders arrive on the same side of the market, and a maker who fills them accumulates inventory in one direction with no opposing informed flow to unwind into. Even a maker who sets out only to capture the spread is left holding a position, not by choice but because one-directional flow gives her nothing to flatten against.[8] She can return to flat only by crossing the spread herself, selling the contract back or buying its complement under Kalshi's netting rule, which on a one-cent spread recovers almost nothing. At the same time, holding the residual is cheap relative to equity trading, because the contract is a bounded claim worth \$0 or \$1 at a known and usually near date and the maker has already posted collateral equal to her maximum loss. Nor is there a continuously traded underlying against which to hedge.

The maker's economics thus come to resemble underwriting more than intermediation, a characterization that concurrent work on Kalshi's NFL markets reaches directly [Palumbo, 2026]. The on-chain record of Polymarket confirms the behavior. Using the data of Akey, Grégoire, Harvie, and Martineau [2026], we find in unreported results that the median maker position is carried to settlement in full, and that 92% of positions in markets resolving on a single event (the closest analog to our single-name markets) are held to resolution; holding persists even in markets whose flow was balanced enough for makers to flatten into. That makers hold even when they could flatten indicates holding is not merely the residue of one-directional flow but is itself profitable in expectation, a possibility to which the behavioral surplus we document in Section 5 gives content. The data thus point to makers overwhelmingly carrying positions to settlement, and we measure maker profitability

[8] An equity-market analog to this situation is merger arbitrage: when insiders know a deal announcement is imminent and buy the target's securities, the dealer supplying the stock is left short into the announcement and cannot flatten, because the informed flow runs one way.

on that basis. We do not, however, rule out that they also capture spread where two-sided flow allows, which is why the break-even condition we develop next retains a spread term alongside the loss on positions held to resolution.

3.2.3 Maker Viability

Under this model of market making, losing on individual trades is structural rather than pathological. Binary settlement guarantees that a maker who sells YES at a fair price of 80 cents loses four times in five, and her edge comes not from any single trade but from the composition of order flow. Writing α for the share of flow that is informed, ℓ for the maker's loss on an informed contract (the full distance from the price to settlement), σ for the surplus she earns on a quasi-informed contract, and s for her spread income per contract, she provides liquidity only if expected profit is non-negative,

$$\underbrace{s + (1-\alpha)\,\sigma}_{\text{spread plus behavioral surplus}} \quad \geq \quad \underbrace{\alpha\,\ell}_{\text{adverse-selection loss}} \quad .$$

The surplus arises because quasi-informed takers are YES-optimistic on events that settle NO; whether it is large enough, and reliable enough, to satisfy this condition where adverse selection is severe is the question the rest of the paper takes up.

This condition organizes what follows. More severe adverse selection raises the right-hand side, through a higher α or a larger ℓ, and liquidity survives only if the left-hand side rises with it. If the behavioral surplus σ is fixed, the only margin that can adjust is the spread s, so trading costs must climb as adverse selection worsens. This is the canonical prediction, and Section 4 shows that under binary settlement the spread required to satisfy the condition is several times wider than makers post or than competition would sustain. If, instead, the surplus scales with adverse selection, so that markets exposed to more informed trading also generate more behavioral overpayment, the condition can hold at a low and roughly constant spread, and trading costs stay low even where adverse selection is severe. The sections that

follow show that the second case is borne out: single-name markets carry both the highest informed price impact (Section 4) and the largest behavioral surplus (Section 5), while their spreads are only modestly wider. The market is in this sense self-reinforcing. The same salient, narrowly defined events that concentrate private information, and so invite informed trading, are the events that draw the most opinionated and YES-optimistic retail demand, and it is that demand that funds the maker who bears the informed flow. Liquidity remains cheap precisely where adverse selection is worst.

3.3 Data

We access all trade, market, and hourly candlestick quote data via Kalshi's REST API. The sample spans the exchange's full operating history from July 2021 through March 11, 2026, and contains over 7 million traded markets organized into nearly 4 million event series. Figure 1 illustrates the exchange's growth: monthly dollar volume rose from under $1 million in mid-2021 to over $4.1 billion by February 2026. Across all markets, we observe 258.7 million trades on roughly 7.4 million markets. For midpoint construction, we use the opening bid and ask of the enclosing hourly candlestick. These valid quotes (where bid > 0, ask < 100, and ask $>$ bid) are available for approximately 7% of trades in the analysis sample.[9] Trades that fall outside candlestick coverage are excluded from all spread-based calculations.

This last feature is why Kalshi, rather than Polymarket, is the natural venue for studying adverse selection. As we discuss in the previous subsection, adverse selection is canonically compensated through the bid-ask spread, so testing whether that mechanism operates requires observed quotes, and Kalshi publishes the quoted bid and ask in its candlestick history, from which we measure effective spreads directly.[10]

[9] Valid opening quotes exist for 35.6 percent of single-name trades (27.5 percent of single-name markets) but only 5.2 percent of broad-based trades, because Kalshi does not retain hourly candlestick history for most short-lived markets and the broad-based universe is dominated by daily and intraday range brackets on equity indexes and cryptocurrencies. Within both types, coverage tilts toward larger markets; the median covered single-name market traded 7,108 contracts against 843 for uncovered markets, and 1,263 against 310 for broad-based. The estimates in Table 4 are substantially unchanged if we trim both types to the overlapping range of market volume, and the fresh-quote premium does not vary with market size.

[10] Polymarket's public record of trades, reconstructed from on-chain fills, contains executed trade prices

3.4 Market Classification

We classify markets by how concentrated private information about their outcomes is likely to be, focusing on the two categories that Internet Appendix B confirms sit at opposite ends of the informed trading spectrum.

Single-name markets reference outcomes specific to a single company or individual: company earnings-call keyword mentions, Tesla delivery targets, Netflix subscriber milestones, CEO departures, IPO outcomes, Super Bowl celebrity attendance, and similar events. Private information about these outcomes is concentrated among a small number of insiders: corporate executives, employees, family members, and their networks.

Broad-based markets reference macroeconomic indicators, market indices, and other outcomes where information is likely to be dispersed across many agents: S&P 500 ranges, Nasdaq 100, Federal Reserve rate decisions, CPI, GDP, unemployment, Bitcoin and Ethereum ranges, Treasury yields, and foreign exchange rates. No individual or small group is likely to possess a decisive private signal about these outcomes.

This classification parallels the single-stock versus index-ETF comparison in equity markets. A corporate insider who knows Tesla's quarterly delivery numbers has a decisive advantage over market makers quoting Tesla contracts; no individual possesses comparable private information about whether the S&P 500 will close above 5,000 tomorrow. Kalshi's own surveillance program classifies markets along the same axis. The insider-risk scoring framework it introduced in 2026 rates new listings on dimensions led by "corporate KPI/events risk" and "outcome concentration risk," the latter higher when an outcome depends on a single individual rather than a decentralized process, which is the single-name versus broad-based distinction arrived at independently by the exchange.[11] Internet Appendix A sets out

only. For both Kalshi and Polymarket, published trades show YES and NO legs that sum to a dollar, which means prevailing spreads also cannot be recovered from the over-round as in Shin [1992]. For these reasons, a spread-based analysis of adverse selection is feasible only on Kalshi. However, as discussed in Section 3.2, we exploit Polymarket's own advantage, that on-chain settlement renders wallet-level positions observable, to examine whether makers in prediction markets are likely to hold positions to settlement. Kalshi, conversely, does not identify accounts across markets.

[11] See "Kalshi Market Integrity Updates: Risk Scoring, Employment Verification, and Whistleblower Pro-

the classification rules for Kalshi markets and the specific contract series assigned to each category. The remaining markets (primarily politics, sports, entertainment, and weather) do not map cleanly onto either category and are excluded from the main analysis.

Table 1 reports summary statistics for the 478,167 single-name and broad-based markets with at least one recorded trade, which constitute the analysis sample for the rest of the paper. Single-name markets account for 15,516 of these markets, 3.0 million trades, and 260 million contracts; broad-based markets account for 462,651 markets, 38.6 million trades, and 4.0 billion contracts. Each trade is for a specified number of contracts, so the trade count and the contract count are analogous to the number of transactions and the total size traded in equity markets, respectively. We therefore report trade characteristics and settlement rates on both an equal-weighted basis (one observation per trade) and a volume-weighted basis (weighted by contracts, which is equivalent to taking the average across individual contracts).

The two weightings tell broadly consistent stories but can diverge because trades vary in size. Takers on single-name markets buy YES 60.2% of the time equal-weighted and 60.9% of the time volume-weighted, compared with 53.2% and 56.3% respectively on broad-based, consistent with the behavioral YES bias we document in Section 5. Settlement rates differ markedly: only 30.9% of single-name markets settle YES equal-weighted (32.5% volume-weighted) versus 38.4% (39.1% volume-weighted) on broad-based. The gap between the YES-buying rate and the YES-settlement rate—roughly 29 percentage points on single-name versus 15 on broad-based—foreshadows the frequency edge that drives maker profitability.

[Insert Table 1 about here]

3.5 Three Examples: What Happens to Makers?

Before turning to our empirical tests, we present three event markets that illustrate how makers can profit on single-name markets despite informed trading, and what happens when

gram," news.kalshi.com (June 9, 2026). The scoring is not exposed in the market data and postdates our sample, so we cite it as corroboration of the classification rather than as an input to it.

they cannot.

Consider first the Super Bowl Bezos market discussed in the Introduction. On February 3, 2026, the Jeff Bezos market (KXSBGUESTS-26-JBEZ) opened at 62 cents and briefly spiked to 95 cents on uninformed YES speculation until the price plummeted to around 40 cents following a surge in NO takers on February 4. The surge in NO trading is consistent with the *Wall Street Journal* reports that Sigma Alpha Epsilon fraternity brothers, tipped off by Bezos' stepson, had traded on the information that Bezos would not attend. In contrast, a similar Super Bowl "appearance" market was also listed for Mark Wahlberg (KXSBGUESTS-26-MWAH), opening on February 3, 2026 at roughly 77 cents. The Wahlberg market held near this level until the day of the game, when the price plummeted toward zero as traders realized Wahlberg would be a no-show.[12] See Figure 2.

Nonetheless, both markets were profitable for makers due to a pronounced YES-bias. In the case of the Wahlberg market, takers were 58.9% YES-biased, paying an average of 29 cents per YES contract for an outcome that never materialized, allowing makers to earn an aggregate of $1.18 million. Even for the Bezos market, despite the evident presence of informed NO takers, makers still earned $29,274 in total (an average of 3.3 cents per contract) because uninformed YES-biased takers (518,000 contracts at an average of 19 cents) more than offset the losses imposed by the NO flow (370,000 contracts). As shown in Figure 3, the early informed trading imposed losses on makers, but the cumulative effect of uninformed YES-biased takers ultimately dominated.

As suggested by the Bezos event market, the viability of market-making in prediction markets with informed trading thus depends on the profits from uninformed traders overwhelming the costs imposed by informed traders. The risk that this cross-subsidy can fail was vividly illustrated with the Netflix Q4 2025 earnings-mention market "Will Netflix

[12]The Bezos and Wahlberg examples illustrate the two dimensions of informed trading. Bezos featured insider trading in which someone actually knew his attendance plans. A simple Google search can turn up that Wahlberg attended the Super Bowl in 2017, 2019, and 2022, and armed with this "information" some traders may have believed he was more likely to attend. These traders were wrong, but their behavior caused the large price surge before more accurate current information entered the market.

say 'Warner Bros'?" (KXEARNINGSMENTIONNFLX-26JUN30-WARN). Approximately three minutes into the call on January 20, 2026, Co-CEO Ted Sarandos mentioned the company's plans to complete the acquisition of "Warner Brothers", prompting a sharp increase in the market price. However, Kalshi's rules required the exact phrase "Warner Bros," and as the interview progressed without those precise words being uttered, sophisticated NO takers recognized that the spoken "Warner Brothers" would not satisfy the market's resolution criteria. Over the next 40 minutes, 210,000 contracts of NO-side volume overwhelmed 71,000 YES-side contracts, driving the YES price from roughly 90 cents to below 40 cents. Makers lost $85,407 during the interview alone; after accounting for modest pre-interview profits, the market's lifetime maker P&L was -$75,754. Figure 4 displays the YES price and cumulative maker P&L during the interview window.

The question our paper addresses is whether the Bezos pattern (maker profitability despite adverse selection) rather than the Netflix pattern (where informed flow overwhelms the uninformed subsidy) holds systematically across the full Kalshi universe.

[Insert Figures 2, 3, and 4 about here]

4 Adverse Selection in Prediction Markets

4.1 Metrics

Measuring adverse selection in prediction markets requires adapting the standard microstructure toolkit. We employ three measures: Kyle's λ, which estimates the permanent price impact of signed order flow; the Glosten and Harris [1988] decomposition, which separates the price impact into permanent (information) and transitory (inventory) components; and the effective spread, which captures the cost of immediacy paid by takers. Each requires some modification for the prediction market setting. In particular, prediction market prices are bounded between 0 and 100 cents, creating heteroskedasticity that distorts linear regressions in levels; quote data are available only at hourly frequency, limiting the precision of midpoint-

based measures; and the absence of a round trip means that the economic interpretation of each measure differs from its equity-market counterpart.

One standard approach, however, is unavailable to us entirely. In equity markets, the Huang and Stoll [1996] decomposition splits the effective spread into a realized spread (the maker's short-term profit) and an adverse selection component (the post-trade midpoint movement in the direction of the taker's trade), estimated at horizons of 5–30 minutes after each trade. Our hourly candlestick data are too coarse for this purpose: over 91% of trades have an identical candlestick midpoint at the 5-minute horizon (because the next observation falls within the same hourly candle), and roughly half are identical at 30 minutes. Consequently, the resulting adverse selection component is mechanically zero for the majority of observations. We therefore rely on Kyle's λ and the Glosten-Harris permanent component, both estimated from trade prices and signed order flow with no dependence on quote midpoints, as our primary measures of informed trading. The effective spread remains useful as a measure of trading costs, even though its decomposition into realized and adverse selection components is infeasible.

4.1.1 Kyle's λ

Kyle's λ captures the impact of signed order flow on price changes: a higher λ is consistent with greater informed trading. In prediction markets, however, the bounded price space ($[0, 100]$ cents) creates a problem: a 1-cent price change near 50 cents has very different information content than a 1-cent change near 5 cents. Following Tsang and Yang [2026], we address this by working in log-odds space.

For each market j, we aggregate trades into consecutive 5-minute intervals indexed by h and compute signed order flow as $\text{SOF}_{j,h} = \sum_{i \in h} q_i \cdot c_i$, where $q_i \in \{+1, -1\}$ is the taker sign for trade i and c_i is the number of contracts in the trade (i.e., the trade's size). Kyle's

λ_j is the slope from the market-level time-series regression:

$$\Delta p^*_{j,h} = \alpha_j + \lambda_j \cdot \text{SOF}_{j,h} + \varepsilon_{j,h} \tag{1}$$

where $\Delta p^*_{j,h}$ denotes the within-interval price change for market j over 5-minute interval h, computed in one of two ways. In the raw specification, $\Delta p^*_{j,h} = p^{\text{close}}_{j,h} - p^{\text{open}}_{j,h}$, the difference in yes_prices (in cents). In the log-odds specification, $\Delta p^*_{j,h} = \log\big(p^{\text{close}}_{j,h}/(1-p^{\text{close}}_{j,h})\big) - \log\big(p^{\text{open}}_{j,h}/(1-p^{\text{open}}_{j,h})\big)$, where prices are expressed as probabilities (cents divided by 100). The log-odds transformation maps prices from the bounded interval $(0,1)$ to $(-\infty,+\infty)$, so that equal absolute changes correspond to equal proportional changes in the odds ratio regardless of the price level. A higher λ_j means that a given quantity of net order flow moves prices more, the signature of informed trading. We require a minimum of 50 five-minute intervals with trades per market. Markets with insufficient trading activity are excluded from all Kyle's λ and Glosten-Harris estimates.

4.1.2 Glosten-Harris Decomposition

The Glosten-Harris decomposition separates price impact into a permanent component (the portion that reflects new information and is not reversed) and a transitory component (the portion that captures inventory or behavioral effects). Following Glosten and Harris [1988], for each market j, we estimate:

$$\Delta p^*_{j,h} = \gamma_{0,j} + \gamma_{1,j} \cdot \text{SOF}_{j,h} + \gamma_{2,j} \cdot \Delta\text{SOF}_{j,h} + \varepsilon_{j,h} \tag{2}$$

where $\Delta p^*_{j,h}$ is defined in two alternative specifications as in equation (1) and $\Delta\text{SOF}_{j,h} = \text{SOF}_{j,h} - \text{SOF}_{j,h-1}$. The permanent component $\gamma_{1,j}$ is the coefficient on the level of order flow; the transitory component $\gamma_{2,j}$ is the coefficient on its first difference.

The Glosten-Harris specification nests the Kyle regression (it simply adds ΔSOF as a second regressor), so $\gamma_{1,j}$ and λ_j are mechanically related. If SOF and ΔSOF are uncorrelated,

$\gamma_{1,j} = \lambda_j$ exactly; in our data, the cross-market correlation between the two estimates is 0.86 in log-odds space, and the mean R^2 improvement from adding ΔSOF is only 0.0196. We report the Glosten-Harris results not as an independent test of informed trading but for what the transitory component reveals.

The answer is striking. Across the pooled sample of single-name and broad-based markets, the transitory component $\gamma_{2,j}$ is economically negligible (mean ≈ 0) and statistically significant (at the 5% significance level) in only 7% of markets.[13] Virtually all of the price impact from order flow in prediction markets is *permanent*, consistent with information incorporation rather than short-lived inventory or liquidity effects. In equity markets, by contrast, the transitory component is typically large, reflecting market makers' active inventory management and mean-reverting quote adjustments [Huang and Stoll, 1996]. The near-absence of a transitory component here is consistent with the holding behavior we document in Section 3.2: prediction market makers carry directional positions to settlement rather than managing inventory dynamically, so there is no inventory effect to reverse. Given the negligible transitory component, we use the permanent component $\gamma_{1,j}$ as our Glosten-Harris measure of informed price impact in the cross-sectional analysis that follows.

4.1.3 Effective Spread

The effective half-spread measures the cost of immediacy: how much the taker pays above (or below) the prevailing midpoint. For each trade i on market j with taker side $q_i \in \{+1, -1\}$ (YES $= +1$, NO $= -1$) and midpoint M_i computed from the enclosing hourly candlestick bid-ask data:

$$\mathrm{ES}_i = q_i \cdot (P_i - M_i) \tag{3}$$

[13] The results are quantitatively similar across market types: the transitory component is significant in 6% of single-name markets and 8% of broad-based markets. The cross-market correlation between λ_j and $\gamma_{1,j}$ is higher on single-name (0.94) than broad-based (0.77), consistent with a larger share of price impact reflecting permanent information incorporation on markets where insiders are most likely to trade.

where P_i is the `yes_price` of the trade. We compute midpoints from the opening bid and ask of the enclosing hourly candlestick; trades that fall outside candlestick coverage are excluded from all spread-based calculations.

4.2 Results

4.2.1 Cross-Sectional Regressions

We begin with a simple question: do markets where private information is likely concentrated among a small number of insiders exhibit more informed trading than markets where information is widely dispersed? Our main specification regresses market-level adverse selection measures on an indicator for single-name markets:

$$\text{AS}_j = \alpha + \beta_1 \cdot \text{SINGLE\_NAME}_j + \beta_2' X_j + \varepsilon_j \tag{4}$$

where AS_j is one of five market-level adverse selection measures reported in each column of Table 2: Kyle's λ_j in raw prices or log-odds, the Glosten-Harris permanent component $\gamma_{1,j}$ in raw prices or log-odds, or the market-level mean effective half-spread, averaged across all trades matched to an hourly candlestick quote. The vector X_j includes log volume, price level, price level squared, and implied volatility ($\sqrt{p(1-p)}$) for market j. Standard errors are heteroskedasticity-robust (HC1). If information concentration is associated with greater informed trading, $\beta_1 > 0$.

[Insert Table 2 about here]

Table 2 reports the results.[14] Kyle's λ is significantly higher for single-name markets in both raw-price ($\beta = 0.0010$, $t = 9.20$) and log-odds ($\beta = 0.0001$, $t = 10.88$) specifications (columns 1 and 2). The Glosten-Harris permanent component tells the same story ($t = 7.46$

[14]The observation counts differ across columns because each measure imposes different data requirements. Kyle's λ and the Glosten-Harris decomposition require at least 50 five-minute trading intervals per market, reducing the sample substantially; most low-activity markets have too few intervals for reliable estimation. The effective spread requires only a valid candlestick midpoint, so column (5) retains more markets.

and $t = 9.58$ in columns 3 and 4). As noted in Section 4, these estimates are not independent of Kyle's λ ($\rho = 0.86$), but they confirm that virtually all price impact is permanent. Column (5) reports the effective spread, which unlike columns (1)–(4) requires a candlestick midpoint. The single-name coefficient is negative and insignificant ($\beta = -0.16$, $t = -0.84$) in the full sample, but as we show below, this result is sensitive to midpoint staleness: when restricted to the freshest available quotes, the single-name spread premium is positive and significant.

The price impact results are exactly what microstructure models would predict. Markets referencing outcomes known to a handful of insiders exhibit more informed trading, and that information is incorporated permanently into prices. Our main analysis focuses on single-name and broad-based markets because the cross-sectional variation in information concentration is sharpest at these two extremes. But Kalshi lists millions of additional markets across politics, sports, entertainment, weather, and other categories. In Internet Appendix B, we extend the Kyle λ analysis to the full universe of non-parlay Kalshi markets and find that the single-name/broad-based dichotomy does in fact capture the two poles of informed trading: single-name markets exhibit the highest λ of any category, and broad-based markets among the lowest, with politics, sports, entertainment, and other categories falling in between.[15]

4.2.2 Robustness

Table 3 subjects the main finding to several alternative specifications, all using the log-odds Kyle's λ. The single-name effect survives each one. It is highly significant without any controls ($t = 14.39$, column 1) and remains so as we progressively restrict the sample to more actively traded markets: $t = 8.83$ at ≥ 100 intervals (column 2), $t = 8.04$ at ≥ 200 (column 3), and $t = 6.01$ at ≥ 500 (column 4). The coefficient attenuates as the sample narrows (from

[15]The one apparent exception is weather markets, whose overall λ exceeds that of single-name markets. A directional decomposition that separates the price impact of order flow on the correct side of the settlement outcome from flow on the wrong side reveals that this anomaly is driven by uninformed flow that moves prices but is wrong about the terminal value. On the dimension that matters for informed trading, single-name markets dominate every other category.

0.000138 to 0.000039), but it is significant at the 1% level in every specification. The result is not driven by thinly traded markets.

[Insert Table 3 about here]

A separate concern is midpoint quality. Our effective spread measure relies on the opening bid and ask of the enclosing hourly candlestick, so a trade at minute :55 of the hour is matched to a quote observed 55 minutes earlier. If this staleness attenuates our spread estimates, it could explain the insignificant result in column (5) of Table 2. To assess this, we restrict the sample to trades occurring within τ minutes of the opening quote and re-estimate the cross-sectional spread regression.

[Insert Table 4 about here]

Table 4 reports the results. At the freshest midpoints ($\tau = 5$ minutes), the single-name spread premium is positive and significant in both equal-weighted ($\beta = 0.73$, $t = 2.81$) and volume-weighted ($\beta = 1.23$, $t = 4.73$) specifications. As the window expands and the midpoint grows staler, the OLS coefficient declines and turns insignificant, but the volume-weighted estimate remains significant through $\tau = 15$ minutes ($t > 4$ in all three narrow windows). The pattern suggests that single-name spreads are indeed wider, but that midpoint staleness attenuates the estimate in the full sample. The declining sample size (from 28,987 markets at the full window to 17,943 at $\tau = 5$ minutes) warrants caution, though the volume-weighted results, which give more weight to actively traded markets where spread estimates are more reliable, are consistent across all but the widest window.

4.2.3 The Limits of Spread Compensation

Wider spreads on single-name markets are the conventional prediction of adverse selection theory, and in equity markets the wider spread also compensates the maker. Because the loss from an informed trade is typically a small multiple of the spread (a few cents of adverse price movement on a stock that moves in small increments), the markup earned

on uninformed trades can, in the canonical Glosten and Milgrom [1985] equilibrium, cover the losses on informed ones. Under binary settlement the loss side of that balance changes scale. A prediction market maker who sells YES at 57 cents to an informed buyer who knows the event will occur loses $100 - 57 = 43$ cents, the full distance to settlement, while her half-spread contributes the same 2 cents it would in equities.

Section 3.2 wrote the maker's break-even as her spread income plus the behavioral surplus, set against her losses to the informed. Here we quantify the first term and show that, on its own, it cannot bear those losses. Because the informed crowd onto one side of the market, a maker who seeks to profit solely from the spread, scalping in and out and aiming to end each contract flat, is instead filled repeatedly against one-directional flow and accumulates an inventory the informed correctly expect to settle against her. With a fraction α of takers informed, consider a maker selling YES, compensated by the spread alone, who faces informed takers buying YES on a market they know will settle YES. She breaks even on a contract priced at P with half-spread s only if

$$s \geq \alpha \cdot (100 - P), \tag{5}$$

where $100 - P$ is her loss per contract at settlement. At $P = 50$ with $\alpha = 0.20$, the required half-spread is 10 cents, a 20-cent full spread. Empirical effective half-spreads in our data average approximately 3 cents, sufficient to offset at most $s/(100-P) = 3/50 = 6\%$ informed takers, and less than that once the spread must also cover inventory carrying costs and order processing. The same bound, with loss P, applies to a maker buying YES from informed takers on the NO side of a market they know will settle NO, and at extreme prices the constraint tightens further, because the 0–100 cent pricing grid caps the half-spread at the smaller distance to the boundary, $\min(P, 100 - P)$, while the loss on the adverse side is the larger, $\max(P, 100-P)$; at $P = 90$ the widest half-spread the grid allows is 10 cents against a 90-cent loss when the informed are on the NO side, absorbing at most $10/90 \approx 11\%$ informed

takers however aggressively the maker prices.

Spreads on Kalshi do appear to reflect adverse selection risk, in that they are wider on single-name markets at fresh midpoints, where informed trading is more severe. But observed spreads would need to be more than three times wider to break even at even moderate levels of informed trading. The second term in the break-even condition, the behavioral surplus, must do that work, which Section 5 measures directly.

5 Maker Profitability

The analysis above shows that the spread, the first term in the maker's break-even condition, cannot protect prediction market makers the way it protects equity dealers. We now measure the second term, computing maker profits directly from settlement outcomes.

5.1 Computing Maker Profits

Because we observe the taker side and settlement outcome for every trade, we can compute the exact maker profit on each transaction:

$$\pi_i^{\text{maker}} = \begin{cases} (P_i - V_j) \cdot c_i & \text{if taker\_side} = \text{YES (maker is short YES)} \\ (V_j - P_i) \cdot c_i & \text{if taker\_side} = \text{NO (maker is long YES)} \end{cases} \tag{6}$$

where P_i is the yes_price (in cents), $V_j \in \{0, 100\}$ is the settlement value, and c_i is the number of contracts. The computation asks: for every trade, what did the non-aggressing side earn at settlement?[16] This measure assumes that the maker holds each position to settlement rather than offsetting it through subsequent trading, for the reasons set out in Section 3.2.

[16] Our analysis ignores trading fees imposed by Kalshi. Kalshi charges taker fees on all markets, ranging from 0.07¢ to 1.75¢ per contract depending on the price level and market type. Starting in April 2025, makers on some market types also face fees ranging from 0.02¢ to 0.44¢ per contract. Given the complexity and time-varying nature of the fee schedule over our sample period, we do not incorporate fees in our calculations. To the extent that fees reduce maker profits, our estimates overstate true profitability for both market types; the cross-sectional comparison between single-name and broad-based markets is unaffected unless fee structures differ systematically across the two groups.

Our per-contract cents measure also preserves the zero-sum accounting that is fundamental to a two-sided market: every cent the maker earns, the taker loses, and vice versa.[17]

Even if some makers offset positions before settlement, the measure remains informative for cross-sectional comparisons. Consider a maker who sells a single contract for YES at 60¢ on a market that settles NO. Our computation credits a 60-cent profit regardless of whether she held the position or sold it earlier at 40¢ (realizing only 20 cents). These measurement errors attenuate toward zero across our large sample, and the cross-sectional comparison between single-name and broad-based markets is valid as long as offsetting behavior does not differ systematically between the two groups.[18]

5.2 Maker Profits by Market Type

Table 5 reports maker profitability for the 475,646 settled single-name and broad-based markets with trading activity.[19]

[Insert Table 5 about here]

Makers earn far more on single-name markets—the very markets where adverse selection is most severe. Per-contract profits average 1.91 cents on single-name versus 0.82 cents on broad-based, more than twice as much. The maker win rate is 63.4% on single-name but only 53.9% on broad-based. In dollar terms, single-name makers earn $305.89 per market

[17]Bürgi, Deng, and Whelan [2026] report that both makers and takers earn negative average *percentage* returns on Kalshi. This does not violate the zero-sum constraint; it reflects the favorite-longshot bias in an equal-weighted average across price levels. A longshot at 5¢ that loses costs −100%; a favorite at 95¢ that wins returns only +5.3%. Because two-thirds of Kalshi markets trade below 10¢ or above 90¢, the massive percentage losses on longshots drag the average return negative for both sides. Our cents-per-contract decomposition avoids this artifact and makes the cross-subsidy from uninformed to informed trading directly visible.

[18]Active inventory management is more likely on broad-based markets, where correlated markets (e.g., adjacent S&P 500 range brackets) could serve as partial hedges; if anything, this *overstates* broad-based maker profits relative to the hold-to-settlement benchmark, biasing against our finding that single-name makers earn more per contract. The on-chain Polymarket record bounds the magnitude. Across 97,766 crypto price markets, the closest analog to our broad-based group, makers carried 95% of gross volume to settlement over January to March 2026, and 82% even in the most balanced markets, where offsetting flow was readily available.

[19]Maker profit requires a settlement outcome, so unsettled markets are excluded. The 475,646 markets here (14,988 single-name + 460,658 broad-based) represent the subset of Table 1's universe that has both settled and recorded at least one trade.

compared to $69.89 for broad-based, and 66.5% of single-name markets are profitable for the maker versus only 49.4% of broad-based. This is difficult to reconcile with the standard view that adverse selection erodes maker profitability.

5.3 The Frequency-Magnitude Decomposition

Table 5 already hints at the answer. On single-name markets, the maker wins on 63.4% of contracts but earns only 16.34 cents per winning contract while losing 23.13 cents per losing contract. Wins are frequent but small; losses are rarer but larger. Whether the maker is profitable overall depends on whether the win frequency is high enough to overcome the win-loss magnitude gap. To formalize this tradeoff, we decompose expected maker profits into a frequency component and a magnitude component:

$$\begin{aligned} E[\pi^{\text{maker}}] &= \Pr(\text{win}) \cdot W - \Pr(\text{loss}) \cdot L \\ &= \underbrace{(\text{WR} - 0.5)(W + L)}_{\text{frequency edge}} + \underbrace{0.5(W - L)}_{\text{magnitude edge}} \end{aligned} \tag{7}$$

where WR is the win rate, W is the average winning amount, and L is the average loss magnitude (both positive). The decomposition benchmarks the maker against a 50/50 baseline. The frequency edge measures the gain from winning more often than a coin flip would predict: each extra win beyond 50% is worth $W + L$ to the maker, because it both adds W to her ledger and avoids a loss of L she would otherwise have taken. The magnitude edge measures the gain from asymmetric stakes: even at a 50% win rate, the maker would earn $0.5(W - L)$ per contract if winning amounts exceed losing amounts. Table 6 reports the two components.

[Insert Table 6 about here]

On single-name markets, the frequency edge is +5.31 cents and the magnitude edge is −3.40 cents, summing to the net per-contract profit of 1.91 cents. The maker wins often

enough to more than cover the magnitude deficit, but the magnitude deficit is real and substantial: over a third of the frequency edge is consumed absorbing it. Broad-based markets show the same pattern at smaller scale: a +2.01 cent frequency edge absorbs a −1.19 cent magnitude deficit for a net of 0.82 cents. The maker's profitability on either market type rests on winning more often than she loses, often enough to compensate for losing more per loss than she earns per win.

[Insert Figure 5 about here]

5.4 Calibration and the Source of the Frequency Edge

Why do single-name makers win on so many contracts? The answer is a pervasive miscalibration between the implied probabilities encoded in yes_prices and the realized settlement rates. Figure 6 documents this miscalibration with an estimator that weights each market equally. For every settled market we compute the volume-weighted average yes_price that takers paid over the first 80% of the market's trading life, a window that closes well before prices mechanically converge toward the settlement value, and we require at least twenty trades in the window so that the average reflects a functioning market.[20] Each market then contributes one unit of weight to the curve, allocated across ten-cent price buckets in proportion to its own volume within the window, so the figure describes the typical market rather than the largest ones. Perfect calibration lies on the 45-degree line, and the vertical bars give 95% confidence intervals with standard errors clustered by market, since each market's settlement outcome is a single draw.

[Insert Figure 6 about here]

On single-name markets the calibration curve bows below the 45-degree line across the entire probability range. In every price bucket the average price paid for YES exceeds the

[20] A market's trading life runs from its first trade to the moment cumulative volume reaches 99% of its total. The results are insensitive to where the window closes. The single-name gap is +4.4 percentage points when the window is shortened to the first 70% of trading life and +4.5 at 80%; extending it over the full life shrinks the gap to +2.3, precisely because late trades occur at prices that have already converged toward the settlement value. The broad-based gap is +0.2 in both restricted windows.

rate at which the underlying events settle YES, by a margin that is statistically significant throughout and widest in the middle of the range, where contracts priced in the forties settle YES only 37% of the time, a shortfall of 8.6 percentage points. Averaged across markets, the typical single-name market trades 4.5 percentage points above its realized settlement rate (SE 0.44, 6,386 markets).[21] Takers on single-name markets exhibit not simply a classical favorite-longshot bias but a pervasive YES optimism, most pronounced where outcomes are most contested. Broad-based markets, by contrast, track the diagonal closely, with a per-market gap of just 0.2 percentage points (SE 0.11, 114,802 markets) and only the mild signature of the textbook favorite-longshot pattern, YES slightly overpriced at longshot prices and slightly underpriced at favorites.

This miscalibration is the behavioral foundation for the frequency edge. The maker wins in exactly two cells of the taker-side $\times$ settlement table: when the taker buys YES and the market settles NO, or when the taker buys NO and the market settles YES. Using the volume-weighted marginals from Table 1, the predicted win rate for single-name makers under independence between taker direction and settlement would be $0.609 \times 0.675 + 0.391 \times 0.325 = 53.8\%$. Yet the actual win rate is 63.4%, a gap of 9.6 percentage points. Because YES is overpriced at essentially every yes_price on single-name markets, YES-biased takers land on the losing side of settlement more often than their marginal YES-bias alone would predict. On broad-based markets, where takers are close to calibrated, the predicted win rate is $0.563 \times 0.609 + 0.437 \times 0.391 = 51.4\%$ versus an actual rate of 53.9%, for an analogous gap of just 2.5 percentage points.

[21] Independent reporting corroborates this miscalibration. Examining more than 35,000 Kalshi mention markets, the *Wall Street Journal* found that YES bets priced at an implied probability of 50% settled YES only about 40% of the time, with the average YES bettor losing roughly 11% of stake [Mehta, Long, and Ostroff, 2026], a pattern that matches the mid-range of Figure 6, where contracts priced in the forties settle YES 37% of the time.

5.5 Cross-Subsidy Across Settlement Outcomes

The calibration gap produces maker profits only to the extent that it translates into realized settlement outcomes the maker benefits from. Because we observe both the taker side and the settlement outcome for every trade, we can decompose total maker P&L into four cells of the taker-side × settlement table. Table 7 reports the results.

[Insert Table 7 about here]

The maker's entire profit comes from markets that settle NO. On single-name markets, NO-settling markets generate $6.25 million in maker profits, and 78% of these markets are profitable for the maker. YES-settling markets lose $1.66 million, with only 41% profitable. Broad-based markets display the same asymmetry: +$64.5 million on NO-settling versus −$32.3 million on YES-settling. On YES-settling markets, the maker has no analogous frequency advantage. She quotes against flow that turns out to be correct, and the magnitude deficit documented above is no longer cross-subsidized by a positive frequency edge.

Panel B decomposes each settlement outcome by taker side. On NO-settling single-name markets, YES takers generate +$18.22 million in maker profits (397% of net maker P&L), while NO takers impose losses of $11.97 million (-$261% of net). On YES-settling single-name markets, the pattern reverses: YES takers cost the maker $8.31 million (-$181% of net), while NO takers contribute +$6.65 million (+145%). Broad-based markets display the same four-cell structure at larger absolute magnitudes but with tighter margins between the winning and losing sides.

Panel C of Table 7 makes the cross-subsidy structure explicit. On each settlement outcome, the maker collects a behavioral tax from uninformed takers who bet against the realized outcome and pays out an adverse taker flow to those correctly aligned with settlement. On NO-settling single-name markets, YES-biased takers pay a YES-optimism tax of +$18.22 million; adverse taker flow from NO-aligned takers costs the maker $11.97 million. The ratio is 1.52×: the behavioral tax more than covers the adverse flow, and the residual is

the maker's net gain on NO-settling markets. On YES-settling single-name markets, the parallel structure applies in reverse. NO-side takers pay makers a contrarian-NO tax of +$6.65 million, but adverse taker flow from YES-aligned takers costs the makers $8.31 million. The ratio is 0.80×: the behavioral tax falls short of covering adverse flow, and the maker loses on YES-settling markets. Broad-based markets show the same pattern (1.28× on NO-settling, 0.87× on YES-settling), at much smaller magnitudes.

Two forces drive the difference between the NO-settling ratios and the YES-settling ratios. First, as noted above, uninformed takers are YES-biased and systematically bet that events will occur when they will not. The behavioral tax on NO-settling markets is therefore larger in volume than the analogous tax on YES-settling markets. Second, more markets settle NO (67.5% of single-name contract volume), so the NO-settling ledger dominates the maker's overall exposure. The combination of these factors means that overall maker profitability depends on both YES-biased takers and the empirical fact that markets (especially single-name markets) settle NO.

The cross-subsidy on NO-settling markets is the maker's lifeline, but it is also precisely the channel through which informed trading inflicts damage. Informed traders who have private information that a market will settle NO can load up on NO at the prevailing yes_price, and their large correct trades risk swelling the adverse taker flow cell and shrinking the cross-subsidy ratio toward 1. The two examples of Section 3.5 sit on either side of this margin: the informed NO flow in the Bezos market was a small erosion of the 1.52× ratio, while on the Netflix "Warner Bros" market the informed flow overwhelmed the uninformed subsidy, the ratio collapsed, and the maker took a lifetime loss.

5.6 The Evolution of the Cross-Subsidy

The analysis to this point treats the maker's profit function as a snapshot. A natural question is whether the cross-subsidy we document is persistent or a transient feature of a still-maturing market. A common prior would be that as prediction markets attract more

sophisticated participants and better pricing tools, calibration should improve and the behavioral surplus should shrink. Figure 7 evaluates this conjecture by plotting the key components of the decomposition month by month from November 2023 through early 2026.[22]

Panel A plots the *anti-informedness gap* on each market type: the difference between the actual maker win rate and the predicted win rate under independence between taker direction and settlement. The predicted win rate is the win rate for makers implied by the aggregate YES-taker share and the aggregate YES-settlement rate in the absence of any correlation between the two. For example, suppose 60% of takers buy YES and 70% of markets settle NO. Under independence, the maker wins whenever the taker buys YES and the market settles NO (probability $0.60 \times 0.70 = 0.42$) or the taker buys NO and the market settles YES (probability $0.40 \times 0.30 = 0.12$). The predicted win rate is $0.42 + 0.12 = 0.54$, or 54%. If the maker's actual win rate is 64%, the anti-informedness gap is +10 percentage points, meaning taker direction is systematically misaligned with settlement beyond what the marginals alone would produce. A positive gap means takers, in aggregate, are on the losing side more often than random. A negative gap means they are on the winning side more often than random. A gap near zero is consistent with calibrated flow.

[Insert Figure 7 about here]

On broad-based markets, the gap oscillates in a narrow band between −5 and +5 percentage points throughout the sample, consistent with aggregate flow that is close to calibrated. On single-name markets, the pattern is markedly different. The gap was close to zero through 2023 and 2024 but turned consistently positive beginning in late 2024 and grew steadily thereafter, reaching +10 to +18 percentage points by early 2026. Panel B decomposes single-name maker profits into the frequency and magnitude components from

[22]Single-name markets did not exist on Kalshi in meaningful numbers before November 2023. A single market (TikTok ban) traded from April through September 2023, joined by one additional market (Google antitrust) in October 2023, with combined monthly volume of fewer than 5,000 contracts. In November 2023, the category expanded to 20 actively traded markets as Kalshi introduced earnings-mention and other company-specific event series. We begin the time series at this point. Broad-based markets (macroeconomic indicators, crypto, and equity indices) have been listed since Kalshi's launch in July 2021.

equation (7), together with the net mean P&L per contract. The frequency edge rose from near zero in early 2024 to approximately +8 cents per contract by early 2026, while the magnitude deficit widened to roughly −6 cents. Yet the net per-contract profit (black line) remained anchored in a narrow one- to three-cent band throughout, indicating that the two components grew in tandem rather than pulling apart.

Both the divergence between single-name and broad-based markets in Panel A and the widening of the frequency and magnitude components in Panel B appear to commence in early 2025, coinciding with the dramatic growth in trading volume documented in Figure 1. We do not attribute the finding causally to the volume increase, as the trend continues monotonically rather than jumping and leveling off. But its direction is clear: the growth in Kalshi trading has deepened rather than diluted the behavioral miscalibration that sustains the cross-subsidy.

5.7 Why the Miscalibration Persists

A miscalibration this large, this systematic, and this durable invites an obvious objection. A trader who recognized that single-name YES prices exceed realized settlement rates could simply sell YES on every single-name market, collect the overpricing, and let competition among such traders drive the gap away. Table 8 shows why this does not happen. Because we observe the taker side of every trade, we can decompose the realized maker book into the two one-sided strategies such a trader could run. The offer-only strategy takes the maker side only when a taker buys YES, so it is short YES on every fill, and the bid-only strategy takes the maker side only when a taker buys NO, so it is long YES on every fill. The two strategies partition the full book, and their P&L sums to the maker profits documented in Table 7.[23]

[23]Like equation (6), the comparison assumes positions are held to settlement. The assumption is conservative here, since allowing early exit would not rescue the one-sided strategy; as discussed below, single-name price impact is almost entirely permanent, so exits after informed flow arrives occur at prices that already impound the news. The Polymarket evidence in Section 3.2 indicates that makers overwhelmingly hold positions to resolution even when two-way flow would permit flattening.

[Insert Table 8 about here]

Panel A confirms that the overpricing is exploitable in expectation. Selling YES alone earns $9.91 million gross across the 14,988 settled single-name markets, more than twice the $4.58 million the full book earns, and its profits on NO-settling markets are exactly the optimism-tax cell of Table 7. Panel B shows the price of that concentration. The offer-only strategy doubles the mean profit per market but carries more than three times the dispersion, so its risk-adjusted return falls a third short of the full book's. The reason lies in the tail. Every market in the offer-only strategy's worst two P&L deciles settled YES. By refusing to bid, the seller concentrates her exposure precisely on the markets where informed YES takers arrived knowing the event would occur. The bid-only strategy is the mirror image, with its worst deciles settling NO without exception. Only the full book escapes this selection, because the maker's YES bids accumulate inexpensive long positions from wrong-side NO takers, and those positions pay off in exactly the states where her short YES positions are run over. The offer side is the revenue engine and the bid side is the hedge, and neither works alone.

Nor can the one-sided trader manage this tail the way an equity arbitrageur would. Kalshi's netting rule makes exit administratively costless, since offsetting a position releases the original collateral, but it does not make exit economically available. Unwinding a short YES position requires a counterparty selling YES, and the moments when the seller most needs to exit are precisely the moments when flow is one-sided in the other direction. Even when a counterparty appears, the price impact documented in Section 4 is almost entirely permanent, so an exit executed after informed flow arrives occurs at prices that already impound the news and crystallizes most of the loss it is meant to avoid. In the meantime the position is fully collateralized and cannot be levered, so the capital that funds the trade is committed until a round trip or settlement. These are the frictions of Shleifer and Vishny [1997] in a stark form. Correcting the bias requires holding the unpopular side of a binary event against exactly the informed-flow risk that makes the position dangerous. Nor is there

a pricing remedy: in a competitive limit order book no participant holds the market power with which the bookmakers of Levitt [2004] and Shin [1991] shade their odds (Section 2), so the only way to trade against the bias is to absorb the flow it generates.

The miscalibration therefore persists not because it is unexploitable but because the binding constraints channel its exploitation through two-sided market making. The traders who harvest the optimism tax are the makers themselves, earning the modest, diversified per-contract profits documented in Table 5 rather than the larger but fragile returns of one-sided arbitrage. This is also why the gap can widen as the market grows, as Figure 7 shows it has. Growth brings optimistic takers faster than it brings capital willing to hold binary risk to settlement, and the makers who absorb their flow have no mechanism, and no incentive, to price the bias away.

6 Flow Toxicity: VPIN in Prediction Markets

Section 3.2 asked whether a maker can gauge the toxicity of order flow in real time. The Volume-Synchronized Probability of Informed Trading (VPIN), developed by Easley, López de Prado, and O'Hara [2012], measures the degree to which order flow is one-sided within equal-sized volume buckets and has been shown to identify elevated toxicity in the hours preceding the May 2010 Flash Crash [Easley, López de Prado, and O'Hara, 2011].[24] Adapting VPIN to prediction markets requires several modifications, but the resulting metric successfully identifies when informed trading is eroding the maker's profits on single-name markets.

What this adds is twofold. The profit decomposition of Section 5 is *ex post*, telling the maker only after settlement how a market's flow treated her, whereas VPIN is *ex ante*,

[24]Prior studies conjectured that VPIN's predictive power is simply a volatility effect. Easley, López de Prado, O'Hara, and Zhang [2021] refute this claim, showing in their extensive futures sample that VPIN and VIX are virtually uncorrelated and perform very differently, particularly out of sample. Easley, Michayluk, O'Hara, Patel, and Putniņš [2026] provide further evidence from equity data noting "the information flows factor [of which VPIN is a component] is not a volatility effect. Whether measured by idiosyncratic volatility, total volatility, or the volatility of share turnover, the information flows beta's effect on cross-sectional asset returns is not affected by these volatility measures, and these volatility measures, in turn, have no significant explanatory power in cross-sectional asset pricing tests".

asking whether the toxicity that erodes her profit is visible in real time, while she can still withdraw. That is the form in which adverse selection actually confronts a liquidity provider. The answer divides along the same single-name versus broad-based line as the rest of the paper, in a way that is itself informative. Trailing toxicity predicts maker losses on single-name markets but not broad-based ones, because one-sided flow signals private information about a terminal outcome in the first case and merely the recent path of a moving underlying in the second.

6.1 Adapting VPIN to Prediction Markets

In conventional equity markets, VPIN partitions trading activity into n volume buckets of equal size V, classifies each trade as buyer- or seller-initiated, and computes:

$$\text{VPIN} = \frac{1}{n} \sum_{\tau=1}^{n} \frac{|V_\tau^B - V_\tau^S|}{V_\tau} \tag{8}$$

where V_τ^B and V_τ^S are the buy and sell volumes in bucket τ. VPIN ranges from 0 (perfectly balanced flow) to 1 (completely one-sided flow).

Three features of Kalshi's market structure require modification of the standard equity VPIN. First, trade classification is exact. The taker side field directly identifies whether the aggressor bought YES or NO, eliminating the classification error inherent in the bulk volume classification or Lee-Ready approaches used in equity markets. Second, volume must be measured in dollars rather than contracts, because a trade of 1,000 contracts at 90¢ deploys nine times the capital of the same trade at 10¢; using contract counts would treat these as equivalent. Third, bucket sizing must accommodate extreme trade-size skew. Across our sample of single-name and broad-based markets, the median trade is 19 contracts, but the 99th percentile is 1,000 and the maximum exceeds 1,000,000 (skewness = 262). The standard VPIN bucket size of average daily volume divided by 50 [Easley, López de Prado, and O'Hara, 2012] produces buckets so small that individual whale trades would fill en-

tire buckets, mechanically pushing VPIN toward 1.0. Indeed, applying the standard rule to our data (dividing each market's average daily contract volume by 50), the median bucket contains just 7 contracts. At this scale, individual trades routinely fill entire buckets with one-sided flow, and mean VPIN exceeds 0.96 across all market types with a median of 1.0, too compressed to discriminate between informed and uninformed flow. We address this by using large, fixed dollar-amount buckets ($500 in our baseline specification) and requiring each market to generate at least five full buckets for inclusion.[25][26]

These modifications also bear on the debate over what VPIN measures. Andersen and Bondarenko [2014] raise three principal criticisms of the standard equity implementation, to which Easley, López de Prado, and O'Hara [2014] reply, and the first two do not arise here by construction. The sensitivity of the metric to bulk-volume versus tick-rule classification is moot, because the taker-side field classifies every trade exactly, and the look-ahead embedded in standardizing VPIN by its full-sample distribution is absent, because the trailing measure we construct below is a raw imbalance average that uses only information available through the prior bucket. Their third criticism, that VPIN's predictive content may reflect volatility rather than information, we address directly in Section 6.3.

A fourth issue is specific to binary settlement and has no equity-market analog. As a market approaches its settlement date and the outcome becomes clear, the price converges toward 100 (for YES-settling markets) or 0 (for NO-settling markets). This convergence mechanically generates one-sided flow as nearly all late trades push in the direction of the terminal value. As such, the resulting VPIN inflation is not a signal of informed trading but instead reflects price convergence that would occur even if every trader were uninformed. This settlement-outcome composition confounds any naive relationship between VPIN and maker

[25] When a trade's dollar volume exceeds the remaining capacity of the current bucket, the trade is split with the portion that fills the current bucket allocated to the current bucket, and the remainder allocated to the next bucket. Buy/sell classification and maker P&L are then allocated proportionally by the fraction of the trade's volume assigned to each bucket. Only complete buckets (those that reach the $500 threshold) are included in the analysis.

[26] We sweep bucket sizes from $100 to $50,000. The results are qualitatively robust. At $500, the sample includes 4,312 single-name and 77,329 broad-based markets. We report the $500 specification as the best balance of sample size and discriminating power.

losses given that high-VPIN buckets will reflect markets approaching a known settlement outcome.[27]

We address this challenge by restricting the analysis to trades executed at interior prices, defined as a volume-weighted average yes_price between 30 and 70 cents within the bucket. At these prices the maker's exposure is economically meaningful, because neither side of the trade is nearly certain to win, and the mechanical link between VPIN and settlement outcome is broken.[28] A market trading at 50 cents is roughly equally likely to settle YES or NO regardless of whether recent flow has been one-sided, so any relationship between VPIN and maker losses at interior prices reflects genuine flow toxicity rather than price convergence.

6.2 Trailing VPIN as a Real-Time Predictor

We construct VPIN as a trailing indicator that a maker could monitor in real time. For each market, we partition pre-close trades into $500 dollar-volume buckets and compute the average absolute imbalance over the most recent 20 buckets:

$$\text{Trailing VPIN}_t = \frac{1}{20} \sum_{\tau=t-20}^{t-1} \frac{|V_\tau^B - V_\tau^S|}{V_\tau} \tag{9}$$

By construction VPIN is direction-agnostic: the absolute value makes it equally high whether one-sided flow runs toward YES or toward NO, so it captures the intensity of order imbalance, not its sign. The question is whether this trailing measure, computed from data available at bucket $t-1$, predicts whether the maker loses money on trades in bucket t. Because every market settles at $1 or $0, we can compute the maker's realized P&L on every bucket and test this directly.

[27] This is not a concern in equity markets, where prices do not converge to a known terminal value. It is a general feature of binary contract markets and would affect any flow-imbalance metric applied to prediction market data without correction.

[28] In the unrestricted sample, the YES-settlement rate among single-name bucket observations rises from 36% in the lowest trailing-VPIN quintile to 95% in the highest. In the interior sample, it ranges from 24% to 30%, confirming that the restriction eliminates the settlement-outcome confound.

The pooled panel contains 32,151 interior-price bucket observations across 534 single-name markets and 544,142 interior-price observations across 17,298 broad-based markets. Table 9 reports maker outcomes by quintile of trailing VPIN.

[Insert Table 9 about here]

On single-name markets, the pattern is consistent with the Section 5 decomposition operating in real time. In the lowest trailing-VPIN quintile (Q1, mean = 0.54), the maker earns an average of \$41 per bucket and loses money on 43.1% of buckets. Profitability is roughly stable through Q4 (mean VPIN = 0.80, P&L = +\$46, loss rate 44.8%). In Q5 (mean VPIN = 0.90), maker P&L turns negative (−\$12) and the loss rate jumps to 51.7%. The YES-settlement rate varies only from 24% to 30% across quintiles, confirming that the interior restriction has eliminated the settlement-outcome confound.

On broad-based markets, trailing VPIN has no predictive power at interior prices. Maker P&L is positive and roughly flat across all five quintiles (ranging from \$4 to \$11), and the loss rate barely moves (49.5% to 49.7%). The absence of predictive power is at first puzzling, given both the Panel A results for single-name markets and VPIN's established track record in equity markets. But the resolution is clear once we consider the nature of the underlying events.

For VPIN to predict maker losses, the *direction* of one-sided flow must predict the *settlement outcome*, because it is the settlement outcome that determines whether the maker wins or loses. On single-name markets, this condition holds. Among buckets in the highest trailing-VPIN quintile (Q5) on single-name markets, those with YES-directed flow settle YES 35.9% of the time, while those with NO-directed flow settle YES only 22.0% of the time ($t = 12.44$, $p < 0.001$).[29] The direction of one-sided flow carries information about whether the event will occur, and that information translates directly into maker losses when the flow

[29]We classify a bucket as YES-directed if the mean signed order flow (taker buy volume minus sell volume, normalized by total volume) over the trailing 20 buckets is positive, and NO-directed otherwise. The test is restricted to interior-price buckets (VWAP 30–70 cents) to eliminate the settlement-outcome confound described above.

is on the correct side.

On broad-based markets, the condition fails. Among Q5 buckets, the YES-settlement rate is 50.0% when flow is YES-directed and 50.2% when flow is NO-directed, statistically indistinguishable ($t = -0.62$, $p = 0.53$).[30] One-sided flow tells the maker nothing about which way the market will settle. The reason is that broad-based markets reference continuously moving underlyings. Consider a Bitcoin daily range market trading at 50 cents. If Bitcoin rises in the preceding hour, takers buy YES, generating one-sided flow and high VPIN. But Bitcoin can just as easily reverse before settlement. The one-sided flow reflects the *recent trajectory* of a continuously moving price, not private information about the *terminal value*. On a single-name market such as "Will Netflix say 'Warner Bros'?", by contrast, one-sided NO flow during the earnings call reflects real-time observation that the phrase has not been uttered—information that is directly and irreversibly about the settlement outcome.[31]

6.3 Regression Analysis

To formalize these patterns, we estimate pooled bucket-level regressions of maker P&L on trailing VPIN, with standard errors clustered by market. Our regression specification takes the form:

$$\pi_{j,t}^{\text{maker}} = \alpha + \beta \cdot \text{Trailing VPIN}_{j,t} + \varepsilon_{j,t} \tag{10}$$

where $\pi_{j,t}^{\text{maker}}$ is the maker's realized dollar P&L on bucket t of market j, and $\text{Trailing VPIN}_{j,t}$ is the mean absolute imbalance over the 20 buckets preceding bucket t, as defined in equation (9). We estimate this specification via OLS and also as a logit with a binary dependent

[30] The predictive capacity of flow direction for single-name (but not broad-based) markets persists in a logit regression of YES-settlement on trailing signed flow among Q5 interior-price buckets, with standard errors clustered by market (single-name: $\hat{\beta} = 0.68$, $z = 2.06$, $p = 0.04$; broad-based: $\hat{\beta} = -0.00$, $z = -0.02$, $p = 0.98$).

[31] We confirm this interpretation by examining the distribution of maker P&L in high-VPIN buckets. On single-name markets, the fraction of buckets with large maker losses (P&L $< -\$100$) rises from 35.7% in Q1 to 47.7% in Q5, while large wins (P&L $> +\$100$) decline slightly (47.3% to 44.5%). High VPIN on single-name markets shifts the maker's outcome distribution toward losses. On broad-based markets, both large wins and large losses increase proportionally in Q5 (47.9% and 47.3%, up from 39.9% and 38.8% in Q1). High VPIN on broad-based markets increases the *variance* of maker outcomes without shifting the *mean*, because one-sided flow is equally likely to be on the winning or losing side of settlement.

variable equal to one if $\pi_{j,t}^{\text{maker}} < 0$ (the maker lost money on the bucket). Table 10 reports the results.

[Insert Table 10 about here]

On single-name markets, both specifications confirm the quintile pattern. Higher trailing VPIN predicts lower dollar P&L ($\hat{\beta} = -139$, $t = -2.42$, $p = 0.016$) and a higher probability that the maker loses on the next bucket (logit $\hat{\beta} = 0.93$, $z = 3.58$, $p < 0.001$).[32] On broad-based markets, neither specification is significant (OLS $t = 0.56$; logit $z = 0.10$), consistent with the absence of concentrated private information in macro and index markets.[33]

Because Section 5 established that the maker's profit is concentrated on NO-settling markets, we can sharpen the test by estimating the regression separately within each settlement outcome. VPIN enters each subsample identically and non-directionally; only the realized settlement outcome, which determines the maker's P&L, differs. On NO-settling single-name markets, where the maker's optimism-tax profits are at risk from informed NO flow, trailing VPIN significantly predicts maker losses: OLS $\hat{\beta} = -160$ ($t = -1.99$, $p = 0.047$). On YES-settling single-name markets, the coefficient is economically negligible and statistically insignificant ($t = -0.09$). The pattern is consistent with the mechanism identified in Section 5: informed traders erode maker profits specifically on NO-settling markets by buying

[32]This relationship is not a volatility effect of the kind Andersen and Bondarenko [2014] attribute to equity-market VPIN. Re-estimating the single-name regression with controls for within-bucket realized volatility, under a price-dispersion definition (the volume-weighted dispersion of transaction prices within the bucket) and a return-based definition (bucket-to-bucket log-odds changes), entered either contemporaneously or as trailing averages over the same twenty buckets as the VPIN measure, leaves the coefficient essentially unchanged, with the OLS estimate between -135 and -157 (t between -2.10 and -2.44) and the logit z between 3.19 and 3.60. The volume dimension of their critique is fixed across observations by the dollar-bucket construction itself, since every bucket contains \$500 of volume by design.

[33]The findings are not sensitive to where the interior band is drawn. Re-estimated on a wider 20–80 cent band, the single-name coefficient is -132 ($t = -2.87$; 660 contracts), and on a narrower 40–60 cent band it is -114 ($t = -1.47$; loss logit $z = 5.9$; 420 contracts), so the sign and magnitude are stable across bands, and the narrowest band, which discards most of the sample, retains significance in the loss-probability specification though not in the dollar specification. The broad-based coefficient is insignificant at every band (OLS t between 0.4 and 0.8), and the collapse of maker profits in the highest trailing-VPIN quintile appears at all three bands. The comparison also confirms the choice of 30–70 as the featured window; within the 20–80 band the YES-settlement rate rises 18.5 percentage points from the lowest to the highest trailing-VPIN quintile, so part of the settlement-convergence confound returns as the band widens, whereas within 30–70 the spread is six points.

NO correctly, and trailing VPIN detects this erosion in real time.

Sorting single-name bucket observations into finer bins reinforces the finding that the toxicity signal is concentrated at the extreme. Makers earn positive P&L at every level of trailing VPIN below approximately 0.86, but at the highest levels (VPIN above 0.94) maker P&L turns sharply negative (−$68 per bucket, loss rate 59.4%). The maker's real-time risk from informed flow materializes not as a gradual erosion across the VPIN distribution but as a discrete jump at the upper tail.

7 Discussion

The preceding sections document an equilibrium in which makers survive severe adverse selection without the spread that canonically compensates it. This section draws out what that equilibrium implies, first for how microstructure economics should model these markets, and then for the regulation of informed trading in them.

7.1 Implications for Market Microstructure Modelling

The behavioral surplus we document has no counterpart in canonical microstructure models. In equity markets the marginal trader is typically a sophisticated institutional participant, and the Glosten and Milgrom [1985] zero-profit condition binds through the spread, the maker's losses to informed traders balanced by her gains from liquidity traders rebalancing portfolio holdings. In prediction markets, neither side of this equilibrium has a direct equivalent. The maker typically takes and holds a directional position rather than earning a spread on a round trip, and there are no liquidity traders in the Glosten-Milgrom sense because the instrument serves no hedging or portfolio-rebalancing function; the uninformed takers here are instead the quasi-informed of Section 3.2, opinionated but without a decisive signal. What matters for the maker's profit is accordingly not the informed-versus-uninformed split but which side of settlement a taker turns out to be on, a distinction no feasible spread can

accommodate, as the six-percent calculation in Section 4 makes precise.

What replaces the spread is the cross-subsidy between these two subpopulations of takers, the second revenue term in the maker's break-even condition from Section 3.2, and with it the maker's economics become those of an insurer rather than a dealer. She collects a premium from miscalibrated takers who buy YES on markets that mostly settle NO, loses on the bad draws (i.e., the informed takers who are on the correct side of settlement), and writes enough premium across the pool to profit in the aggregate. Because the surplus is concentrated in the single-name markets where adverse selection is most severe, the premium is largest precisely where the insurance is most needed, which is why makers bear the heaviest informed flow at a spread no wider than the one on broad-based markets.

The equilibrium has a precise structural form, and Section 5 characterized it in two complementary ways. At the contract level, the frequency-magnitude decomposition determines the maker's per-contract profit as a frequency edge (the gain from winning on more contracts than the marginals alone would produce) offset by a magnitude edge (the net of win and loss sizes). At the settlement-outcome level, the cross-subsidy of behavioral tax against adverse taker flow determines the maker's P&L on each outcome. Both views capture the same equilibrium. On single-name markets, the ratio of behavioral tax to adverse taker flow is 1.52× on NO-settling markets and 0.80× on YES-settling markets, with the NO-settling excess dominating the YES-settling shortfall, so the maker earns 1.91 cents per contract net. Unlike the Glosten-Milgrom spread, which the maker sets to equate expected profit to zero, the ratio in this equilibrium is not chosen by any participant. It emerges from the joint distribution of taker direction and market settlement.

The time series evidence in Section 5 speaks to the stability of this equilibrium in a way the cross-section alone cannot. As Kalshi scaled from 0.3 million to more than 80 million single-name contracts per month, both the behavioral tax and the adverse taker flow grew in tandem. The frequency edge rose from near zero to approximately eight cents per contract, while the magnitude deficit widened to roughly six cents. Yet the net per-contract profit

remained stable in a narrow one- to three-cent band throughout. This invariance, with gross components widening at scale while the residual stays pinned near a constant, reveals that the equilibrium has so far been robust in a way that is surprising relative to Glosten-Milgrom intuitions. In aggregate, makers continue to supply liquidity and to earn a stable positive return even on the markets where adverse selection is most severe. But while our data establish this equilibrium in the current Kalshi market structure, we cannot predict its persistence going forward.

Formalizing this equilibrium is a natural direction for future theoretical work. Open questions include characterizing the conditions under which the ratio remains above one, the speed at which it adjusts to shocks, and the welfare incidence across its subpopulations. Our contribution here is to document the equilibrium empirically, measure its magnitudes, and establish its stability over the period in which prediction markets have scaled.

7.2 Implications for Informed Trading Regulation

The equilibrium also reshapes the market-quality case for regulating insider trading here. In the Glosten-Milgrom framework, informed trading is what allows the market to achieve efficiency by incorporating new information. But it also widens the spread and deters liquidity, which reduces market quality for all participants. This latter effect justifies enforcement of insider trading laws on welfare grounds, because enforcement restores the maker's incentive to quote. In the equilibrium we describe, correct-side takers reduce the behavioral-tax-to-adverse-flow ratio, but while that ratio exceeds one the maker still supplies liquidity at a profit, and we observe no liquidity collapse over the 2.5 years single-name trading has scaled. The market-quality rationale thus does not bind here as it does in equity markets.

Market quality, though, is only one rationale for policing insider trading. The large number of bills now circulating in Congress seeking to impose new insider trading rules in prediction markets illustrates how broader concerns about fairness and the deterrence of oppor-

tunistic trading on nonpublic information loom over these markets as they grow.[34] Certainly, our own finding that a behavioral tax is paid by YES-biased takers who both lose to makers and finance the rents the informed extract is hard to reconcile with a well-functioning market.

At the same time, restricting trading on nonpublic information carries a well-known cost. In conventional securities markets, a parity-of-information regime, which would bar all such trading, has long been resisted for the burden it places on price discovery. It is for this reason that, in fashioning insider-trading liability under Rule 10b-5, the Supreme Court conditioned liability on the breach of a duty, an approach the Commodity Futures Trading Commission has largely followed in Rule 180.1.[35] Under that standard, trading on misappropriated information, in breach of a duty to its source, is prohibited; trading on an analytical edge is not. The Court drew that line deliberately because a parity-of-information rule would chill the research that makes prices accurate. But it also means this logic imposes no liability where the subject of a market trades on his own knowledge of the outcome, since one cannot misappropriate from oneself, or where he voluntarily shares that knowledge with others, as may have occurred in the opening Bezos example.

The novel policy question is thus whether this same duty-versus-possession line should be drawn in prediction markets as it has been in conventional securities and derivatives markets. Indeed, Kalshi itself imposes on its members a rule consistent with a parity-of-information regime, prohibiting trades by anyone with access to nonpublic information about a contract's subject, a rule that in principle would capture the Bezos Super Bowl trades.[36] Polymarket and

[34]See, e.g., the Prediction Markets Security and Integrity Act of 2026, S. 4060; the Prediction Market Act of 2026 (Gillibrand–McCormick); the Public Integrity in Financial Prediction Markets Act of 2026, H.R. 7004; the Public Integrity in Financial Prediction Markets Act, S. 4188; the PREDICT Act, H.R. 8076; and the End Prediction Market Corruption Act, S. 4017, all 119th Cong. (2026). Several bills that would instead bar categories of event contracts cite insider-trading concerns as a rationale, among them the BETS OFF Act (S. 4115, H.R. 7955) and the STOP Corrupt Bets Act of 2026 (S. 4226, H.R. 8123). These bills regulate trading conduct, and are distinct from the Commission's unexercised authority to bar an event contract by subject matter as contrary to the public interest, for instance one concerning terrorism, assassination, war, or gaming. See 17 C.F.R. §40.11.

[35]*Chiarella v. United States*, 445 U.S. 222 (1980); *Dirks v. SEC*, 463 U.S. 646 (1983); see CFTC Rule 180.1, 17 C.F.R. §180.1; Cong. Rsch. Serv., LSB11406, *Prediction Markets and Insider Trading Law* (2026).

[36]Kalshi Rulebook Rule 5.17(y)–(z). This prohibition has been in Kalshi's rulebook since its early versions

its CFTC-approved exchange, by contrast, largely track the conventional misappropriation theory, substantially reducing the likelihood that such trades would have been prohibited.[37] Prediction markets have in this way reopened the parity-versus-duty question that proved decisive in shaping insider-trading liability in securities markets.[38]

We take no position on which approach is right; reasonable minds differ, as they did a generation ago. What our findings add is the stakes on each side. A parity rule would chill informed trading on the correct side of settlement, the very flow that our directional estimates show pulling single-name prices toward their true values. A duty standard preserves that price discovery but leaves the fairness cost in place, as uninformed takers continue to pay the optimism tax that finances the rents the informed extract.

8 Conclusion

Adverse selection is the central challenge for any market that aggregates information through trading. Our paper asks how this challenge plays out in prediction markets, and what happens to the people who provide the liquidity that makes aggregation possible.

Using 41.6 million trades across 478,167 Kalshi markets, we show that the standard tools of equity market microstructure (Kyle's λ, the Glosten-Harris decomposition, effective spreads) detect adverse selection in prediction markets precisely where theory predicts it.

(formerly Rule 5.13) and thus predates essentially all of the trading volume in our sample; what has changed more recently was the addition in February 2026 of an enforcement and surveillance apparatus. Consistent with that timing, the adverse selection we measure is robust within the sample: single-name informed price impact is significantly elevated before the surveillance launch (in 2025, the last full pre-launch year, single-name Kyle's λ exceeds the broad-based benchmark with $t = 15.2$), and the behavioral cross-subsidy is present in every year and grows rather than shrinks (Section 5), the opposite of what progressively tightening enforcement would produce.

[37] Polymarket US (QCX LLC) Rulebook (2026), Rule (g)–(i). The pending bills divide along the same axis. S. 4060 would bar the use of material nonpublic information outright, a possession standard, whereas S. 4017 would direct the Commission to reach only its use in breach of an express or implied duty.

[38] Somewhat surprisingly, none of these approaches to regulating insider trading in prediction markets has adopted the cabined middle ground securities law developed for an analogous problem in the tender-offer setting. Under SEC Rule 14e-3, 17 C.F.R. §240.14e-3, a person with material nonpublic information about a tender offer obtained from the offeror, the target, or their agents may not trade the target's securities, without proof of any duty. An analogous rule could bar trading in a single-event market on information obtained from the subject of the market. See *United States v. O'Hagan*, 521 U.S. 642, 666 (1997) (sustaining Rule 14e-3).

Kyle's λ in log-odds space is significantly elevated on single-name markets ($t = 10.88$), the Glosten-Harris decomposition confirms that this price impact is almost entirely permanent, and effective spreads appear wider at fresh midpoints. But because makers hold their positions to settlement, the link between wider spreads and maker compensation that is central to equity microstructure does not survive. The informed edge on a held position, up to the full distance between the trade price and settlement, can exceed the spread by an order of magnitude, so the conventional spread-compensation mechanism breaks down.

The resolution lies in the maker's profit function. Single-name makers earn 1.91 cents per contract, more than twice the 0.82 cents earned on broad-based markets, despite facing more adverse selection. The profit comes from a frequency edge (+5.31 cents per contract) that outweighs a magnitude deficit (−3.40 cents), which arises from the fact that makers win on 63.4% of contracts, but when they lose, they lose more per contract than they earn per win. The frequency edge is funded by a pervasive YES overpricing on single-name markets—takers pay yes_prices that systematically exceed realized settlement rates at every probability level—and is concentrated precisely where the adverse selection is most severe.

A decomposition by settlement outcome sharpens the picture. The maker's entire profit comes from markets that settle NO, where a YES-optimism tax from uninformed YES takers exceeds offsetting adverse taker flow by 1.52× on single-name and 1.28× on broad-based. Adapting VPIN to prediction markets as a real-time toxicity metric, we show that trailing VPIN at interior prices (30–70 cents, where a settlement-outcome endogeneity specific to binary markets is absent) predicts maker losses on the next volume bucket for single-name markets but not for broad-based. The effect is concentrated on NO-settling markets, consistent with informed traders eroding the maker's cross-subsidy in real time.

The broader implication is that adverse selection and market viability can coexist and even reinforce each other through behavioral cross-subsidization. In effect, the maker's profit is a residual arising from the behavioral tax paid to her by uninformed YES takers on NO-settling markets minus the losses that adverse taker flow imposes. As long as the ratio

exceeds one, a maker earns profits by supplying liquidity. This mechanism, which has no close equivalent in equity markets, may explain why prediction market liquidity has proven more resilient than standard microstructure theory would predict, given the level of adverse selection we document. It also suggests that the health of these markets depends less on the severity of informed trading than on the continued willingness of uninformed takers to participate.

References

Akey, P., Grégoire, V., Harvie, N., and Martineau, C. (2026). Who wins and who loses in prediction markets? Evidence from Polymarket. Working paper, ESSEC Business School and University of Toronto.

Ali, M. M. (1977). Probability and utility estimates for racetrack bettors. *Journal of Political Economy*, 85(4), 803–815.

Andersen, T. G., and Bondarenko, O. (2014). VPIN and the flash crash. *Journal of Financial Markets*, 17, 1–46.

Arrow, K. J. (2008). The promise of prediction markets. *Science*, 320(5878), 877–878.

Bagehot, W. (1971). The only game in town. *Financial Analysts Journal*, 27(2), 12–14, 22.

Becker, J. (2025). Microstructure of wealth transfer in prediction markets. Working paper, `jbecker.dev`.

Bürgi, C., Deng, C., and Whelan, K. (2026). Makers and takers: Economics of Kalshi. CEPR/CESifo Working Paper.

Copeland, T. E. and Galai, D. (1983). Information effects on the bid-ask spread. *Journal of Finance*, 38(5), 1457–1469.

Dowie, J. (1976). On the efficiency and equity of betting markets. *Economica*, 43(170), 139–150.

Easley, D., López de Prado, M. M., and O'Hara, M. (2011). The microstructure of the "Flash Crash": Flow toxicity, liquidity crashes, and the probability of informed trading. *Journal of Portfolio Management*, 37(2), 118–128.

Easley, D., López de Prado, M. M., and O'Hara, M. (2012). Flow toxicity and liquidity in a high-frequency world. *Review of Financial Studies*, 25(5), 1457–1493.

Easley, D., López de Prado, M. M., and O'Hara, M. (2014). VPIN and the flash crash: A rejoinder. *Journal of Financial Markets*, 17, 47–52.

Easley, D., López de Prado, M. M., and O'Hara, M. (2016). Discerning information from trade data. *Journal of Financial Economics*, 120(2), 269–286.

Easley, D., López de Prado, M. M., O'Hara, M., and Zhang, Z. (2021). Microstructure in the machine age. *Review of Financial Studies*, 34(7), 3316–3363.

Easley, D., Michayluk, D., O'Hara, M., Patel, V., and Putniņš, T. (2026). Information flows and asset pricing. *Review of Finance*, forthcoming.

Easley, D., Kiefer, N. M., O'Hara, M., and Paperman, J. B. (1996). Liquidity, information, and infrequently traded stocks. *Journal of Finance*, 51(4), 1405–1436.

Butcher, S. (2025). Electronic market makers on Kalshi are small fish in a big pond. *eFinancialCareers*, `https://www.efinancialcareers.com/news/electronic-market-makers-trading-on-kalshi-are-small-fish-in-a-big-pond`.

Gârleanu, N., Pedersen, L. H., and Poteshman, A. M. (2009). Demand-based option pricing. *Review of Financial Studies*, 22(10), 4259–4299.

Glosten, L. R. and Harris, L. E. (1988). Estimating the components of the bid/ask spread. *Journal of Financial Economics*, 21(1), 123–142.

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), 71–100.

Grossman, S. J. and Miller, M. H. (1988). Liquidity and market structure. *Journal of Finance*, 43(3), 617–633.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance*, 46(1), 179–207.

Hasbrouck, J. (2004). *Empirical Market Microstructure*. Oxford University Press.

Hauser, F. and Rittler, D. (2024). Price formation in field prediction markets. *Journal of Financial Markets*, forthcoming.

Ho, T. and Stoll, H. R. (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics*, 9(1), 47–73.

Huang, R. D. and Stoll, H. R. (1996). Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE. *Journal of Financial Economics*, 41(3), 313–357.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6), 1315–1335.

Levitt, S. D. (2004). Why are gambling markets organised so differently from financial markets? *Economic Journal*, 114(495), 223–246.

Lee, C. M. C. and Ready, M. J. (1991). Inferring trade direction from intraday data. *Journal of Finance*, 46(2), 733–746.

Madhavan, A. (2002). Market microstructure: A practitioner's guide. *Financial Analysts Journal*, 58(5), 28–42.

Mitts, J. and Ofir, M. (2026). From Iran to Taylor Swift: Informed trading in prediction markets. Working paper, Columbia University and University of Haifa.

Ng, J., Peng, L., Tao, R., and Zhou, G. (2026). Price discovery and trading in modern prediction markets. SSRN Working Paper No. 5331995.

Mehta, N., Long, K., and Ostroff, C. (2026). Why almost everyone loses—except a few sharks—on prediction markets. *Wall Street Journal*, May 3, 2026.

O'Hara, M. (1995). *Market Microstructure Theory*. Blackwell, Cambridge, MA.

Palumbo, N. (2026). A microstructure perspective on prediction markets. SSRN Working Paper No. 6325658.

Shin, H. S. (1991). Optimal betting odds against insider traders. *Economic Journal*, 101(408), 1179–1185.

Shin, H. S. (1992). Prices of state contingent claims with insider traders, and the favourite-longshot bias. *Economic Journal*, 102(411), 426–435.

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

Swanson, E. T., Wang, C., and Wu, Y. (2025). Kalshi and the rise of macro markets. Federal Reserve Board Working Paper.

Tsang, K. P. and Yang, J. (2026). Political shocks and price discovery in prediction markets. arXiv:2603.03152.

Wolfers, J. and Zitzewitz, E. (2004). Prediction markets. *Journal of Economic Perspectives*, 18(2), 107–126.

Wolfers, J. and Zitzewitz, E. (2006). Interpreting prediction market prices as probabilities. NBER Working Paper No. 12200.

Figures and Tables

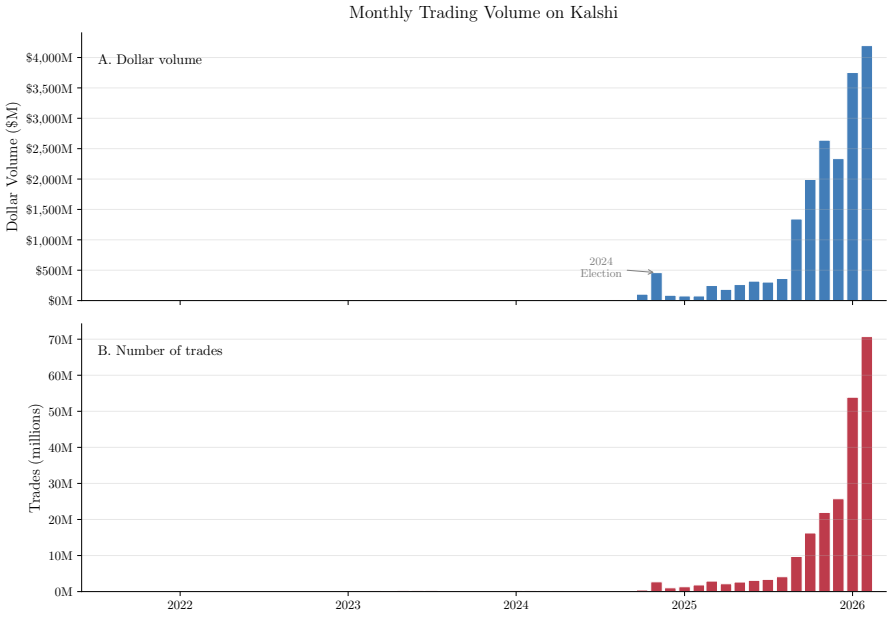

Figure 1: Monthly trading volume on Kalshi. Panel (a) shows dollar volume; Panel (b) shows number of trades. Peak monthly dollar volume exceeded $4.1 billion in February 2026.

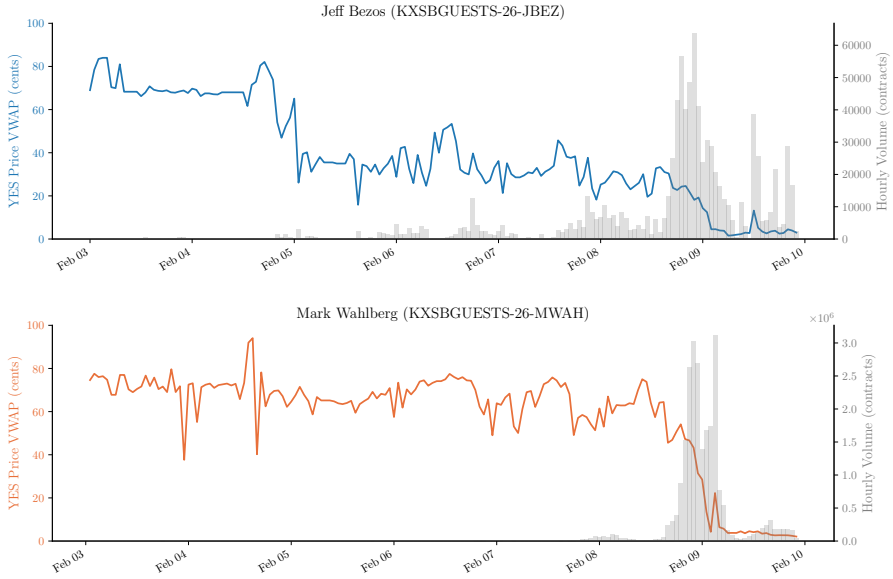

Figure 2: Hourly YES-price VWAP (blue, left axis) and trading volume (gray bars, right axis) for the Super Bowl LX attendance markets on Jeff Bezos (top) and Mark Wahlberg (bottom). Both markets settled NO.

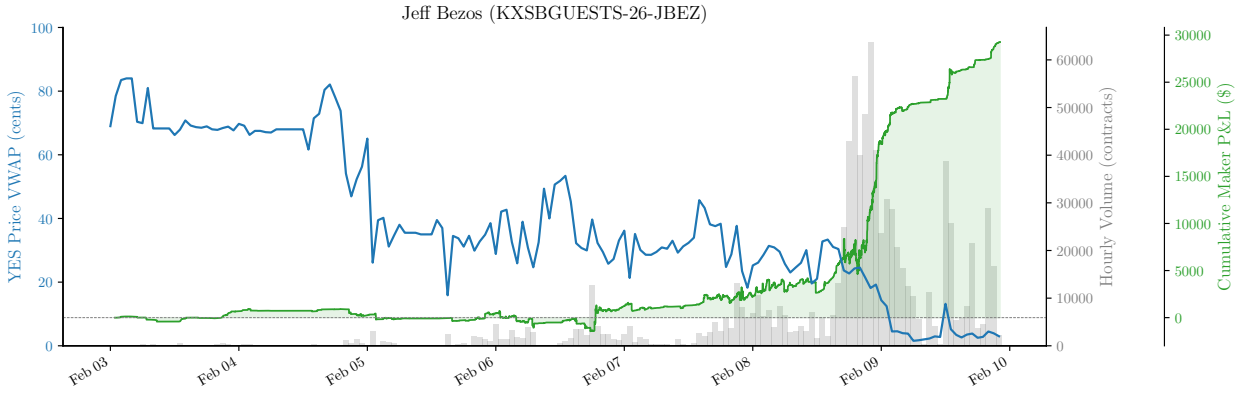

Figure 3: Hourly YES-price VWAP (blue, left axis), trading volume (gray bars, first right axis), and cumulative maker P&L (green, second right axis) for the Bezos Super Bowl market. Final maker P&L was $29,274.

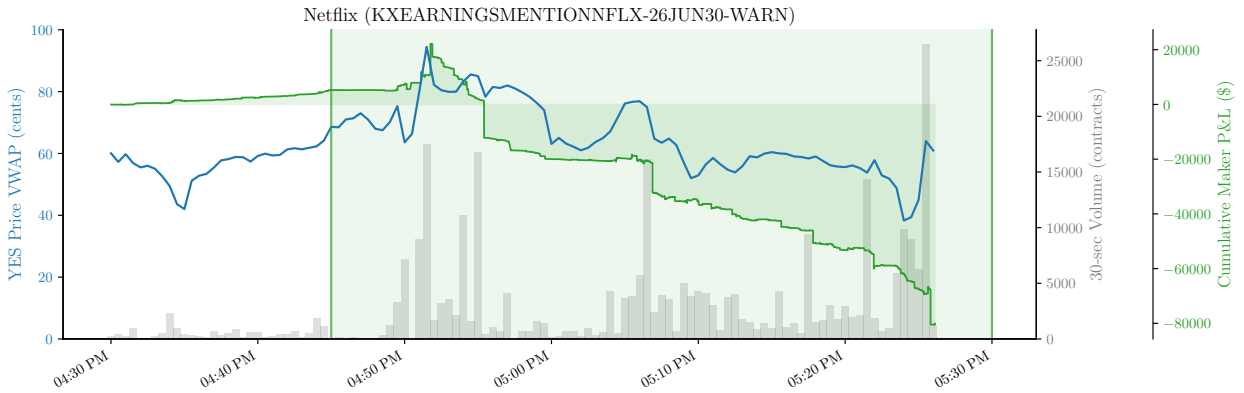

Figure 4: Thirty-second VWAP of the YES price (blue, left axis), trading volume (gray bars, first right axis), and cumulative maker P&L (green, second right axis) for the Netflix "Warner Bros" earnings-mention market during the earnings-day interview window (4:30–5:45 PM ET). Vertical lines mark the interview start and end; the shaded region spans the live broadcast.

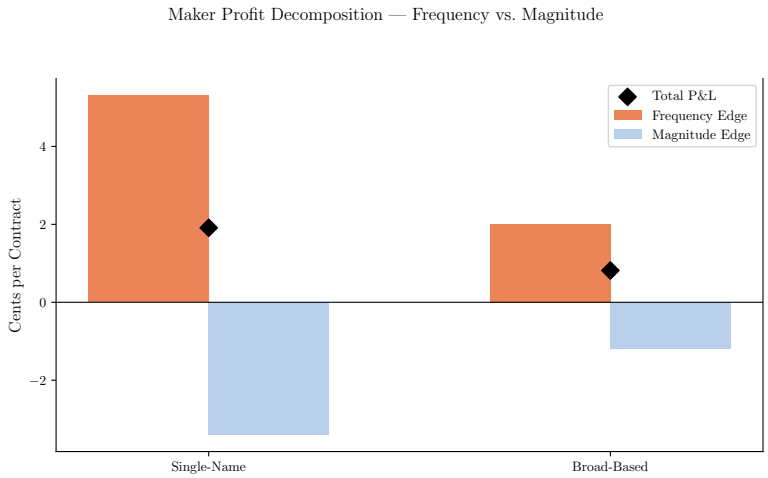

Figure 5: Decomposition of maker profits per contract into the frequency edge and the magnitude deficit, by market type. Diamonds mark total P&L per contract.

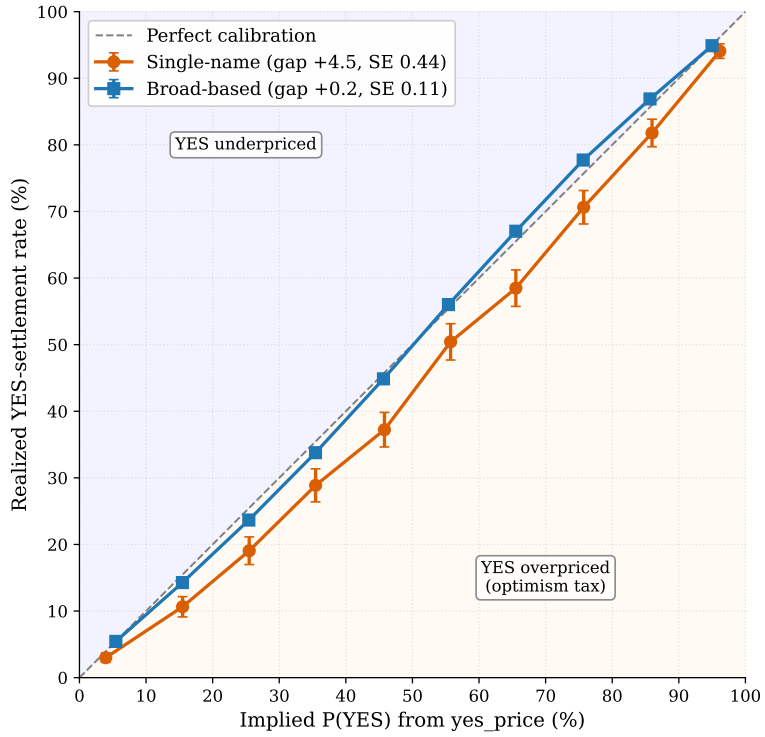

Figure 6: Market calibration. Each settled market with at least twenty trades in the first 80% of its trading life contributes one unit of weight, allocated across ten-cent price buckets in proportion to its own window volume. The horizontal axis gives the weighted mean YES price in each bucket, the vertical axis the realized YES-settlement rate, and the 45-degree line perfect calibration. Vertical bars are 95% confidence intervals with standard errors clustered by market.

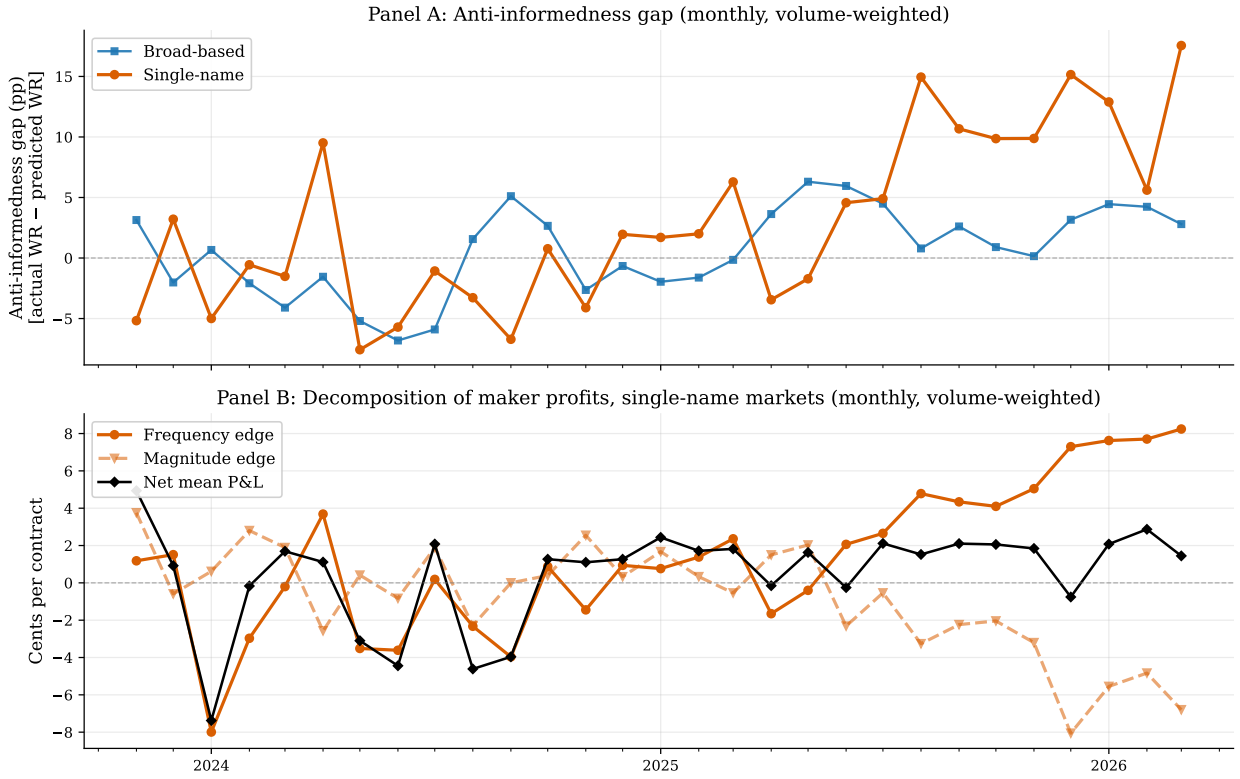

Figure 7: Evolution of the cross-subsidy, monthly (volume-weighted). Panel A plots the anti-informedness gap on each market type, defined as the difference between the actual maker win rate and the predicted win rate under independence between taker direction and settlement. Panel B decomposes single-name maker profits into the frequency edge and magnitude edge from equation (7), together with the net mean P&L per contract.

Table 1: Summary Statistics

	Single-Name	Broad-Based	Total
Markets	15,516	462,651	478,167
Settled markets	14,988	460,658	475,646
Total trades	3,022,734	38,572,352	41,595,086
Total contracts	259,898,094	4,019,457,227	4,279,355,321
Trade characteristics — equal-weighted			
Mean yes_price	42.0¢	45.9¢	
Mean taker price (cost basis)	47.3¢	50.0¢	
Taker buys YES (%)	60.2%	53.2%	
Trade characteristics — volume-weighted			
Mean yes_price	38.1¢	40.3¢	
Mean taker price (cost basis)	38.9¢	46.8¢	
Taker buys YES (%)	60.9%	56.3%	
Settlement — equal-weighted			
Settles YES (%)	30.9%	38.4%	
Settles NO (%)	69.1%	61.6%	
Settlement — volume-weighted			
Settles YES (%)	32.5%	39.1%	
Settles NO (%)	67.5%	60.9%	

Notes: The analysis sample consists of single-name and broad-based markets with at least one recorded trade (see Internet Appendix A for classification). Mean taker price is the taker's average cost basis, equal to yes_price when the taker buys YES and 100−yes_price when the taker buys NO. Volume-weighted trade statistics weight each trade by its number of contracts; volume-weighted settlement rates weight each settled market by its total contract volume.

Table 2: Cross-Sectional Regressions: Adverse Selection on Market Type

	(1) Kyle λ (raw)	(2) Kyle λ (log-odds)	(3) GH perm. (raw)	(4) GH perm. (log-odds)	(5) Eff. spread
SINGLE_NAME	0.000993***	0.000114***	0.000897***	0.000107***	-0.158059
	(9.20)	(10.88)	(7.46)	(9.58)	(-0.84)
Controls	Yes	Yes	Yes	Yes	Yes
N	14,801	14,801	14,504	14,504	28,987
R^2	0.047	0.040	0.039	0.036	0.019

Notes: The unit of observation is the market (ticker); each column reports OLS estimates with HC1 standard errors. The dependent variable is Kyle's λ from 5-minute interval regressions in columns (1)–(2), the permanent component of a Glosten-Harris decomposition in columns (3)–(4), and the market-level mean effective half-spread from trades matched to an hourly candlestick quote in column (5). Controls include log volume, price level, price level squared, and implied volatility. t-statistics in parentheses. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Table 3: Robustness: Kyle's λ (Log-Odds) Under Alternative Specifications

	(1) No controls	(2) ≥ 100 intervals	(3) ≥ 200 intervals	(4) ≥ 500 intervals
SINGLE_NAME	0.000138***	0.000058***	0.000051***	0.000039***
	(14.39)	(8.83)	(8.04)	(6.01)
Controls	No	Yes	Yes	Yes
N	14,801	5,872	2,427	857
R^2	0.020	0.074	0.086	0.130

Notes: The unit of observation is the market (ticker); OLS with HC1 standard errors; the dependent variable is Kyle's λ in log-odds space. Column (1) uses the baseline sample (≥ 50 five-minute intervals per market); columns (2)–(4) progressively restrict to more actively traded markets and include controls (log volume, price level, price level^2, implied volatility). t-statistics in parentheses. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4: Effective Spread and Midpoint Freshness

	OLS (equal-weighted)		WLS (volume-weighted)		
Window (τ)	β (SINGLE_NAME)	t	β (SINGLE_NAME)	t	N
≤ 5 min	+0.731	2.81	+1.225	4.73	17,943
≤ 10 min	+0.371	1.59	+0.995	4.17	21,663
≤ 15 min	+0.270	1.22	+0.896	4.06	23,541
All (≤ 60 min)	−0.158	−0.84	−0.321	−1.64	28,987

Notes: The specification is that of column (5) of Table 2, with the market-level mean effective half-spread computed only from trades within τ minutes of the opening candlestick quote; controls are computed from all trades on the market. OLS gives each market equal weight; WLS weights by number of trades. t-statistics use HC1 standard errors.

Table 5: Maker Profitability by Market Type

	Single-Name	Broad-Based
Markets (settled)	14,988	460,658
Trades	2,785,039	38,119,402
Contracts traded	239,820,053	3,942,900,273
Per-contract		
Mean maker P&L (cents)	1.91	0.82
Maker win rate (%)	63.4%	53.9%
Average win (cents)	16.34	24.21
Average loss (cents)	-23.13	-26.59
Per-market		
% markets maker profitable	66.5%	49.4%
Mean maker P&L/market ($)	$305.89	$69.89
Median maker P&L/market ($)	$4.48	$-0.01
Total maker P&L ($)	$4,584,642	$32,196,483

Notes: Maker P&L computed using realized settlement values. Per-contract estimates reflect volume-weighted trade means. Per-market statistics aggregate all trades within each ticker.

Table 6: Decomposition of Maker Profits

	Single-Name	Broad-Based
Total expected P&L	+1.91¢	+0.82¢
Frequency edge (Win rate − 50%) × (Avg win + Avg loss)	+5.31¢ (+13.4pp × 39.47¢)	+2.01¢ (+3.9pp × 50.80¢)
Magnitude edge 50% × (Avg win − Avg loss)	-3.40¢ (50% × -6.79¢)	-1.19¢ (50% × -2.38¢)

Notes: Expected P&L, win rate, average win, and average loss are from Table 5. "Avg loss" enters the decomposition as a positive magnitude (i.e., |average loss|). Amounts in cents per contract.

Table 7: Maker P&L by Settlement Outcome and Taker Side

	Single-Name		Broad-Based	
	P&L ($M)	% of net	P&L ($M)	% of net
Panel A: By settlement outcome				
Settles NO	+6.25	+136%	+64.46	+200%
% markets profitable	78.0%		54.8%	
Settles YES	−1.66	−36%	−32.27	−100%
% markets profitable	40.8%		40.9%	
Net P&L	**+4.58**		**+32.20**	
Panel B: By settlement outcome × taker side				
Settles NO, taker buys yes	+18.22	+397%	+295.52	+918%
Settles NO, taker buys no	−11.97	−261%	−231.06	−718%
Settles YES, taker buys yes	−8.31	−181%	−251.73	−782%
Settles YES, taker buys no	+6.65	+145%	+219.46	+682%
Panel C: Cross-subsidy decomposition				
NO-settling markets				
YES-optimism tax (YES takers)	+18.22		+295.52	
Adverse taker flow (NO takers)	−11.97		−231.06	
Ratio	1.52×		1.28×	
YES-settling markets				
Contrarian-NO tax (NO takers)	+6.65		+219.46	
Adverse taker flow (YES takers)	−8.31		−251.73	
Ratio	0.80×		0.87×	

Notes: All P&L figures are in millions of dollars. Ratios above 1 in Panel C indicate that the behavioral tax exceeds the adverse taker cost. Percentages in Panels A–B express each cell as a fraction of the market type's net maker P&L; because the net is the small difference of large, offsetting gross flows, a single gross cell can exceed 100% of the net or carry the opposite sign.

Table 8: One-Sided Exploitation of the YES Overpricing: Limits to Arbitrage

Panel A: Aggregate P&L by settlement outcome ($ millions)

Strategy	NO-settling	YES-settling	Total
Offer-only (sell YES)	+18.22	−8.31	+9.91
Bid-only (buy YES)	−11.97	+6.65	−5.32
Full book (both sides)	+6.25	−1.66	+4.58

Panel B: Per-market P&L distribution

Strategy	Mean ($)	Std. dev. ($)	Mean/SD	Worst markets settling YES
Offer-only (sell YES)	+661	35,722	0.019	100%
Bid-only (buy YES)	−355	25,891	−0.014	0%
Full book (both sides)	+306	10,931	0.028	68%

Notes: 14,988 settled single-name markets. P&L assumes positions are held to settlement, as in equation (6), and is gross of fees; the two one-sided strategies partition the full book, so their rows sum to the full-book row in each column of Panel A. In Panel B each observation is one market; Mean/SD is a Sharpe-style ratio. The final column reports the share of markets in the strategy's bottom two P&L deciles that settled YES.

Table 9: Trailing VPIN and Next-Bucket Maker Outcomes (Interior Prices)

	Mean VPIN	Mean P&L	Median P&L	% Lost	% YES	n
Panel A: Single-Name Markets (32,151 buckets, 534 markets)						
Q1 (low)	0.54	+$41	+$74	43.1%	23.7%	6,431
Q2	0.67	+$36	+$78	44.8%	27.4%	6,430
Q3	0.74	+$42	+$91	44.7%	29.8%	6,430
Q4	0.80	+$46	+$99	44.8%	30.3%	6,430
Q5 (high)	0.90	−$12	−$44	51.7%	29.8%	6,430
Panel B: Broad-Based Markets (544,142 buckets, 17,298 markets)						
Q1 (low)	0.49	+$4	+$4	49.5%	51.5%	108,829
Q2	0.61	+$7	+$7	49.3%	51.4%	108,828
Q3	0.70	+$9	+$7	49.5%	50.0%	108,828
Q4	0.80	+$11	+$12	49.3%	47.9%	108,828
Q5 (high)	0.94	+$10	+$14	49.7%	50.1%	108,829

Notes: Each observation is a $500 dollar-volume bucket with volume-weighted average yes_price between 30 and 70 cents. Trailing VPIN is the mean absolute imbalance over the preceding 20 buckets. Mean and median maker P&L are in dollars per bucket. "% Lost" is the fraction of buckets where the maker's realized P&L is negative. "% YES" is the fraction of bucket observations on markets that settle YES. Quintiles are formed within each market type.

Table 10: Trailing VPIN and Next-Bucket Maker Outcomes: Regression Results (Interior Prices)

	OLS: Maker P&L ($)			Logit: Pr(Maker Loses)			
Subsample	$\hat{\beta}$	t	p	$\hat{\beta}$	z	p	N
Panel A: By market type							
Single-name	−139	−2.42	0.016	0.93	3.58	<0.001	32,151
Broad-based	14	0.56	0.577	0.01	0.10	0.919	544,142
Panel B: Single-name, by settlement outcome							
Settles YES	−8	−0.09	0.924	0.36	0.75	0.451	9,067
Settles NO	−160	−1.99	0.047	1.00	2.91	0.004	23,084

Notes: Panel A reports pooled bucket-level regressions of next-bucket maker P&L on trailing VPIN (20-bucket trailing window), restricted to buckets with VWAP between 30 and 70 cents. Standard errors are clustered by market (ticker) for both OLS and logit. Panel B estimates the single-name regression separately within each settlement outcome. The dependent variable is maker P&L on bucket t (OLS, in dollars) or an indicator for maker P&L < 0 (logit).

Internet Appendices

A. Market Classification

We begin with the full universe of over 7.3 million markets listed on Kalshi through early March 2026, encompassing 258.7 million trades across politics, sports, weather, entertainment, macroeconomic indicators, company-specific events, and other categories. We classify each market as single-name, broad-based, or unclassified based on a single economic criterion: how concentrated is private information about the market's outcome?

Single-name markets reference outcomes specific to a single company, individual, or small organization. A small number of corporate insiders, employees, or close associates possess decisive private information: whether Tesla will exceed a delivery target (known to Tesla executives), whether a specific word will be uttered during an earnings call (known to the speaker), or whether a celebrity will attend an event (known to the celebrity and their circle). The information structure is analogous to that of an individual stock.

Broad-based markets reference macroeconomic aggregates, market indices, or asset prices where information is dispersed across thousands of analysts, traders, and policymakers. No individual or small group possesses a decisive private signal. Examples include S&P 500 index levels, CPI releases, Fed funds rate decisions, and cryptocurrency prices. The information structure is analogous to that of an index ETF or a macro futures contract.

Classification is performed using pattern-matching on the series ticker prefix assigned by Kalshi to each market. Applying these rules to the full universe yields 15,516 single-name markets and 462,651 broad-based markets with at least one recorded trade, which together constitute the analysis sample. The remaining markets (primarily politics, sports, weather, and entertainment) do not map cleanly onto either end of the information-concentration spectrum and are excluded from the main analysis. Politics markets, for example, include both broad-based outcomes that aggregate millions of voter decisions and individual-action

outcomes ("Will John Fetterman vote for RFK Jr.?") where a single decision-maker's intention is private information. Sports markets involve competitive settings where coaches and players have asymmetric information about injuries and strategy, but no individual can unilaterally determine the result. Because these categories blend elements of both concentrated and dispersed information, including them would blur the very distinction our empirical strategy is designed to exploit.

A.1 Single-Name Markets

Table A.1 summarizes the composition of the single-name category. Trump speech content and earnings call mentions together account for the two largest volume subcategories, with earnings mentions spanning over 100 publicly traded companies. These markets are particularly well-suited for studying adverse selection given that (i) the outcome is determined

Table A.1: Single-Name Market Composition

Subcategory	Markets	Series	Volume
Trump speech content	1,306	54	47,890,581
Earnings call mentions	2,635	107	44,151,438
Super Bowl attendance	41	2	40,270,790
Press conference mentions	1,441	3	28,828,987
Spotify rankings	5,642	22	26,388,967
Netflix rankings/subscribers	2,003	12	22,937,244
SpaceX launches/IPO	146	19	9,461,472
IPO timing	345	21	8,395,714
Inauguration attendance	36	3	6,636,894
Company KPIs	251	43	6,331,006
App store rankings	1,434	2	4,855,665
TikTok ban/sale	17	10	4,765,596
M&A outcomes	3	1	2,733,515
CEO departures	13	13	2,524,662
Tech layoffs	3	3	1,783,616
Musk-specific	114	3	1,644,286
OpenAI governance	41	12	1,019,166
Other	45	21	3,598,982
Total	**15,516**	**351**	**264,218,581**

Notes: Volume is in contracts.

during a discrete, observable event; (ii) company insiders and analysts who preview the call have a clear informational advantage; and (iii) each market references a single corporate entity. Other significant subcategories include Spotify daily rankings (where employees at Spotify or music labels with access to real-time streaming data hold an informational edge), Tesla delivery and production targets, Super Bowl celebrity attendance, and various corporate events such as CEO departures, IPO timing, and M&A outcomes.

A.2 Broad-Based Markets

Table A.2 summarizes the composition of the broad-based category. Three subcategories

Table A.2: Broad-Based Market Composition

Subcategory	Markets	Series	Volume
Cryptocurrency (BTC, ETH, etc.)	303,336	116	2,564,137,697
Fed funds / rate decisions	1,198	33	657,890,989
Equity indices (S&P 500, Nasdaq)	86,516	38	569,082,997
Recession / gas prices / debt	605	30	99,801,263
CPI / inflation	2,392	29	61,882,065
Employment / labor	1,091	11	25,293,404
Commodities / rates (oil, gold, Treasuries)	5,112	22	22,335,847
Foreign exchange (USD/JPY, EUR/USD)	61,940	17	19,899,880
GDP	295	26	16,498,151
PCE / other macro	166	4	852,608
Total	**462,651**	**326**	**4,037,674,901**

Notes: Volume is in contracts.

dominate: Fed funds rate decisions, cryptocurrency (Bitcoin, Ethereum, and other digital assets), and equity indices (S&P 500 and Nasdaq 100 range markets). All three share the defining characteristic that information about the outcome is widely dispersed. Specifically, the federal funds rate is set by a 12-member committee based on publicly observed data, cryptocurrency prices are determined by a global decentralized market, and index levels aggregate information across hundreds of constituent stocks.

B. Informed Trading Across the Full Kalshi Universe

The main analysis restricts attention to single-name and broad-based markets, where we hypothesize that the cross-sectional variation in information concentration is sharpest. This appendix extends the analysis to the full universe of Kalshi markets. The goals are to assess whether the single-name/broad-based dichotomy captures the extremes of informed trading or whether other market categories exhibit comparable or greater levels of adverse selection, and to extend the calibration analysis of Section 5 to the same universe. The universe also contains a natural check on our settlement-conditioned measures, a category whose terminal value no trader can hold in advance, and Section B.4 shows that the measures correctly read its one-sided flow as noise rather than information.

B.1 The Kalshi Market Universe

Kalshi lists over 7.3 million markets spanning company-specific events, macroeconomic indicators, politics, sports, entertainment, weather, and other categories. Table B.1 summarizes the universe by market type.

Two features of this table warrant discussion. First, parlay markets account for 86.7% of all listed markets. These are programmatically generated multi-leg markets that bundle two or more standalone propositions into a single bet. Consider a Buffalo Bills game against the Miami Dolphins. Kalshi offers standalone markets on the game's outcome (Will the Bills win?), the total points (Over 47.5?), and individual player statistics (Will Josh Allen throw for over 275.5 yards?). Kalshi also generates every possible combination of these propositions as a separate market: "Bills win AND over 47.5 points" is one ticker; "Bills win AND Josh Allen over 275.5 yards AND Tyreek Hill scores a touchdown" is another. With 10 standalone propositions on a single game, the number of possible multi-leg combinations runs into the hundreds; across a full season of games, the combinatorics produce millions of tickers. These markets are not dormant (95% have volume of at least 10), but they are compositional of the

Table B.1: Kalshi Market Universe

Category	Markets	% of Total	Mean Trades
Single-name	15,516	0.2%	17,029
Broad-based	462,651	6.3%	8,727
Politics	8,400	0.1%	191,119
Sports	340,924	4.6%	97,349
Entertainment	11,612	0.2%	30,339
Weather	45,658	0.6%	12,434
Other	94,357	1.3%	71,190
Parlay (multi-event)	6,385,199	86.7%	633
Total	7,364,317	100%	6,894

Notes: Market counts reflect all markets with at least one recorded trade through March 11, 2026. Mean trades is the average number of trades per market within each category. Single-name and broad-based are defined as in Appendix A. Remaining categories are assigned using series-ticker pattern matching on the full Kalshi taxonomy.

standalone markets that form the basis of our analysis. Because their information structure is inherited from the component legs rather than generated independently, we exclude them throughout.

Second, politics markets are the most actively traded per market (191,119 mean trades), driven by presidential election markets that attracted billions of dollars in volume during the 2024 cycle. Sports markets are next, followed by other and entertainment. Single-name markets, while few in number, are moderately active (17,029 mean trades), providing sufficient data for reliable estimation.

B.2 Kyle's λ Across All Categories

We replicate the Table 2 specification on the full sample of non-parlay markets, with single-name as the omitted category:

$$\hat{\lambda}_j = \beta_0 + \sum_{k \neq \text{SN}} \beta_k \cdot \mathbf{1}[\text{Category}_j = k] + \gamma' X_j + \varepsilon_j \tag{11}$$

where X_j includes the same controls as in the main analysis (log volume, price level, price level squared, and implied volatility). We also report a specification with log market duration as an additional control, since market duration varies substantially across categories (weather markets have a median duration of 39 hours; politics markets, 927 hours). All results include heteroskedasticity-robust standard errors. The sample comprises 129,729 markets with at least 50 five-minute trading intervals.

Table B.2 reports the results.

Table B.2: Kyle's λ (Log-Odds) by Market Category

	(1) Baseline	(2) + log(duration)
BROAD_BASED	-0.000106***	-0.000111***
	(-11.28)	(-11.82)
POLITICS	-0.000073***	-0.000058***
	(-6.78)	(-5.31)
SPORTS	-0.000096***	-0.000102***
	(-10.96)	(-11.68)
ENTERTAINMENT	-0.000117***	-0.000107***
	(-11.15)	(-10.23)
WEATHER	0.000070***	0.000057***
	(7.36)	(5.89)
OTHER	-0.000061***	-0.000065***
	(-6.77)	(-7.20)
LOG(DURATION)		-0.000008***
		(-11.43)
Controls	Yes	Yes
N	129,729	129,729
R^2	0.101	0.102

Notes: OLS with HC1 standard errors; t-statistics in parentheses. Omitted category: single-name (as classified in Appendix A). Multi-event (parlay) markets excluded. Controls: log volume, price level, price level^2, implied volatility. Column (2) adds log(market duration), where duration = close_time − open_time. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Every category exhibits significantly lower Kyle λ than single-name markets, with one exception: weather. Broad-based markets ($\beta = -0.000106$, $t = -11.28$), sports ($t = -10.96$),

entertainment ($t = -11.15$), politics ($t = -6.78$), and other ($t = -6.77$) all have substantially lower price impact of order flow. The magnitudes are also large. For instance, the broad-based coefficient represents a reduction of roughly half relative to the single-name mean. Weather markets, however, exhibit significantly *higher* λ than single-name ($\beta = 0.000070$, $t = 7.36$), and this effect persists after controlling for market duration ($\beta = 0.000057$, $t = 5.89$).

The weather result is puzzling. Who has a decisive private signal about tomorrow's high temperature in Chicago? The next section resolves this puzzle.

B.3 Directional Lambda: Separating Informed from Uninformed Price Impact

Kyle's λ measures the magnitude of price response to order flow, but it does not distinguish whether that order flow is on the correct side of the eventual settlement outcome. In equity markets, this distinction is unavailable because the terminal value is unobserved. In prediction markets, where every market settles at 0 or 100, we can make it directly.

For each market j with settlement value $V_j \in \{0, 100\}$, we classify each 5-minute interval h based on whether signed order flow is on the correct side of settlement:

$$\text{CorrectSide}_{j,h} = \mathbf{1}\big[\text{sign}(\text{SOF}_{j,h}) = \text{sign}(V_j - 50)\big] \tag{12}$$

Net buying (SOF > 0) is correct-side when $V_j = 100$ (the market settles YES), and net selling (SOF < 0) is correct-side when $V_j = 0$. Intervals with zero net flow are excluded. We then estimate λ separately on each subset:

$$\text{Correct-side:} \quad \Delta\ell_{j,h} = \alpha_j^C + \lambda_j^C \cdot \text{SOF}_{j,h} + \varepsilon_{j,h}^C \quad \text{for } h \in \{\text{CorrectSide} = 1\} \tag{13}$$

$$\text{Wrong-side:} \quad \Delta\ell_{j,h} = \alpha_j^W + \lambda_j^W \cdot \text{SOF}_{j,h} + \varepsilon_{j,h}^W \quad \text{for } h \in \{\text{CorrectSide} = 0\} \tag{14}$$

where $\Delta\ell_{j,h} = \log\big(p^{\text{close}}_{j,h}/(1-p^{\text{close}}_{j,h})\big) - \log\big(p^{\text{open}}_{j,h}/(1-p^{\text{open}}_{j,h})\big)$ is the within-interval price change in log-odds space, as in the log-odds specification of equation (1). We require at least 25 intervals per subset. The directional difference $\lambda^D_j = \lambda^C_j - \lambda^W_j$ captures whether order flow has more price impact when it is on the informationally correct side. A positive λ^D is the signature of genuine informed trading: prices move more when the flow knows something.

A design note: we split on the direction of the *regressor* (SOF relative to settlement) rather than on the direction of the *dependent variable* (the price change). Splitting on the dependent variable would introduce selection bias, since intervals with large $|\Delta\ell|$ would be mechanically more likely to be classified as toward-settlement.

Table B.3 reports the decomposition in three panels. Panel A establishes the benchmark on single-name markets directly. Order flow on the correct side of settlement moves prices sharply ($\lambda^C = 0.000567$, $t = 30.64$); flow on the wrong side barely moves them at all ($\lambda^W = 0.000038$, $t = 3.08$); and the gap between the two is large and overwhelmingly positive ($\lambda^C - \lambda^W = 0.000488$, $t = 22.26$). This is the signature of informed trading in its cleanest form: on the markets where private information is most concentrated, prices respond to flow that turns out to be right and all but ignore flow that turns out to be wrong. Panel B then asks how every other category compares, reporting cross-sectional regressions of each component on category dummies with single-name as the omitted group.

Panel B shows that this single-name benchmark is the ceiling, and the decomposition resolves the weather puzzle. Column (1) reports the price impact of order flow on the correct side of settlement. On this dimension, single-name markets dominate every category: broad-based ($t = -20.90$), entertainment ($t = -18.33$), sports ($t = -16.13$), politics ($t = -12.03$), other ($t = -10.12$), and weather ($t = -2.22$) all exhibit significantly lower correct-side λ. When order flow is on the right side of the settlement outcome, single-name markets show the strongest price response.

Column (2) reveals the source of weather's anomalous overall λ: wrong-side price impact. Weather markets have the highest wrong-side λ of any category ($\beta = 0.000103$, $t = 7.50$),

Table B.3: Directional Lambda: Correct-Side vs. Wrong-Side Price Impact

	(1) λ^C (correct-side)	(2) λ^W (wrong-side)	(3) $\lambda^C - \lambda^W$
Panel A: Single-name markets (levels)			
Mean	0.000567***	0.000038***	0.000488***
	(30.64)	(3.08)	(22.26)
N contracts	3,145	3,714	2,806
Panel B: Difference from single-name (category dummies)			
BROAD_BASED	-0.000382***	-0.000004	-0.000354***
	(-20.90)	(-0.27)	(-16.10)
POLITICS	-0.000258***	0.000040***	-0.000237***
	(-12.03)	(2.61)	(-8.91)
SPORTS	-0.000299***	-0.000020*	-0.000211***
	(-16.13)	(-1.66)	(-9.31)
ENTERTAINMENT	-0.000366***	0.000001	-0.000309***
	(-18.33)	(0.07)	(-12.75)
WEATHER	-0.000044**	0.000103***	-0.000121***
	(-2.22)	(7.50)	(-4.94)
OTHER	-0.000195***	-0.000008	-0.000072***
	(-10.12)	(-0.60)	(-2.98)
Controls	Yes	Yes	Yes
N	85,830	98,457	60,131
R^2	0.098	0.015	0.053
Panel C: Within-category levels (unconditional means)			
BROAD_BASED	0.000129***	0.000026***	0.000094***
	(32.36)	(5.77)	(18.44)
POLITICS	0.000148***	0.000070***	0.000058***
	(12.11)	(7.35)	(3.89)
SPORTS	0.000142***	-0.000005***	0.000174***
	(50.95)	(-4.73)	(39.95)
ENTERTAINMENT	0.000122***	0.000038***	0.000069***
	(13.49)	(4.99)	(6.15)
WEATHER	0.000568***	0.000139***	0.000431***
	(68.33)	(23.77)	(40.34)
OTHER	0.000283***	0.000019***	0.000328***
	(41.42)	(6.98)	(29.41)

Notes: t-statistics in parentheses. Panel A reports unconditional means of each component across single-name contracts, with one-sample t-tests (paired for column 3). Panel B reports OLS regressions of each component on category dummies with HC1 standard errors; the omitted category is single-name, so each coefficient is the difference from single-name conditional on the controls below. Controls: log volume, price level, price level^2, implied volatility, log(market duration). Panel C reports the Panel A statistics computed within each remaining category, so column (3) gives each category's own correct-minus-wrong price-impact gap. Column (3) contract counts: broad-based 7,326; politics 1,447; sports 21,978; entertainment 2,284; weather 14,845; other 9,447. Sample restricted to settled non-parlay markets with ≥ 50 total intervals and ≥ 25 intervals per directional subset. Correct-side: $\text{sign}(\text{SOF}) = \text{sign}(V_j - 50)$. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

meaning that prices on weather markets respond more strongly to order flow that turns out to be on the losing side of settlement than on any other type of market. Prices move, but the flow is wrong. Politics also shows a positive wrong-side coefficient ($t = 2.61$), but the effect is smaller. For broad-based, sports, entertainment, and other markets, wrong-side price impact is statistically indistinguishable from single-name.

Column (3) reports the directional difference $\lambda^C - \lambda^W$. Every category has a significantly negative coefficient: the gap between correct-side and wrong-side price impact is largest for single-name markets. This is exactly what informed trading should produce. When insiders trade on single-name markets, prices respond, and when noise traders push prices in the wrong direction, the response is muted. Weather is anomalous in a different way. Its wrong-side impact is the largest of any category even as its correct-side impact falls below single-name, so a uniquely large share of weather's price movement is generated by flow that carries no information about the terminal value. Section B.4 develops the mechanism.

Panel C completes the picture by reporting the same decomposition within each remaining category. The correct-minus-wrong gap is positive and statistically significant everywhere, from broad-based markets ($t = 18.44$) to politics ($t = 3.89$), so flow on the correct side of settlement moves prices more than wrong-side flow in every corner of the exchange, and informed trading is present throughout the universe in diluted form. The Panel C levels are not comparable across categories, because price impact per contract is mechanically larger in thinner markets; cross-category comparison is the role of the controlled regressions in Panel B. Read together, the panels also sharpen the weather puzzle. Within weather, correct-side impact still exceeds wrong-side impact, so weather flow is not systematically backwards at the interval horizon; the anomaly is that so much of weather's price movement, absolutely and relative to every other category, comes from flow that turns out to be wrong.

The overall picture is consistent with the main analysis. The single-name/broad-based dichotomy captures the two extremes of informed trading on Kalshi. Single-name markets exhibit the highest price impact from order flow, and that price impact is concentrated on

the correct side of the settlement outcome. Broad-based markets sit at the opposite end, with the lowest correct-side λ and a negligible directional gap. The remaining categories fall between these poles, with weather providing an instructive case in which high overall λ reflects uninformed price impact rather than genuine information.

B.4 Weather: Correlated Noise in Nearly Pure Form

Weather is the one category whose terminal value no trader can hold in advance, since settlement is a realized measurement like the day's high temperature in Chicago, yet Table B.2 assigns it the highest overall λ of any category and Table B.3 the largest wrong-side price impact. It is therefore a natural check on our measures: a category where one-sided flow demonstrably cannot carry decisive private information, so whatever moves its prices is what a false positive looks like. Three facts reveal the mechanism. First, the wrong-side impact concentrates in thin markets. Within weather, wrong-side λ falls from 0.000233 in the thinnest volume tercile to 0.000060 in the thickest ($t = 9.41$), and it resides in the daily temperature series (e.g., the day's high temperature in Chicago) rather than the category's other markets (e.g., whether it rains in New York). Second, it is not an artifact of contract design. Weather's temperature brackets share their single-winner field structure with the broad-based closing-range fields (e.g., the range in which the S&P 500 will close on a given day), yet the weather fields carry two and a half times the wrong-side impact of the broad-based ones ($t = 8.22$). The structure is the same, so the difference must lie in the underlying. Third, the direction of sustained one-sided weather flow anti-predicts the outcome. Applying the flow-direction test of Section 6 to weather markets, high-imbalance episodes with YES-directed flow settle YES 39.7% of the time while those with NO-directed flow settle YES 65.8% of the time, a difference of -26 percentage points (cluster-robust $z = -3.13$), the mirror image of the strong positive prediction on single-name markets and unlike the null on broad-based markets.

These facts are consistent with what one would expect of weather prediction markets.

Weather traders share the same public signal (i.e., the forecast and the intraday observation of the temperature itself), so their order flow is correlated by construction; a forecast revision moves many traders in the same direction at once. In thin books that common flow moves prices substantially. But the terminal value is a realized measurement whose deviation from the forecast nobody holds in advance, so the flow that moves weather prices carries no private information about settlement, and flow that chases a trend which subsequently reverts is systematically wrong. Weather is thus the case in which one-sided order flow is correlated noise in nearly pure form, and it is detectable as such only because settlement is observed. The same one-sidedness that the directional decomposition assigns to the wrong side on weather markets would register in a conventional, non-directional λ or VPIN as apparent informed trading. This is the identification concern of Section 3.2 made concrete, and it is why the settlement-conditioned measures, rather than the levels of price impact, carry our inferences about information.

B.5 Calibration Across Categories

The preceding sections extend our price-impact measures to the full universe. This section does the same for the calibration analysis of Section 5. Two features of our classification bear on how those categories should be read. First, the single-name group was constructed to isolate the markets most plausibly exposed to informed trading, so that adverse selection could be studied where it should be most acute. Its rule-based definition is accordingly conservative, admitting only series that fit the definition unambiguously, and some markets assigned to politics, entertainment, or the residual group could defensibly be called single-name as well. The clearest example is the family of mention markets, which ask what a named person will say or do at a scheduled event and which the rules capture only under recognized series. Their presence works against our comparisons rather than for them, since it moves the excluded categories toward the single-name benchmark. Second, as Panel C of Table B.3 establishes, informed trading is present throughout the universe in diluted

form. Someone may glimpse a review score before an embargo lifts, read a whip count faster than the market, or learn of an injury before it is announced. The question is therefore not whether the excluded categories contain information or optimism, but how much of each, and in which contracts.

Figure B.1 reports the calibration curve for each category, estimated exactly as in Figure 6, with each settled market that has at least twenty trades in the first 80% of its trading life contributing one unit of weight, allocated across price buckets in proportion to its own window volume, and with standard errors clustered by market. Broad-based markets are essentially calibrated (a per-market gap of +0.2 percentage points), and sports (+1.3), weather (+1.0), and politics (+1.1, with a wide standard error of 0.58) sit close to the diagonal. Entertainment is the clear exception, with a gap of +3.1 that approaches the single-name benchmark of +4.5, and the residual group is also elevated (+2.4), partly reflecting the mention-type markets assigned to it; treating those markets as single-name moves the residual group toward the diagonal and widens the single-name gap. Measured over full trading histories, then, the optimism tax is concentrated where our design placed it, with entertainment as the one category that shares it to a meaningful degree.

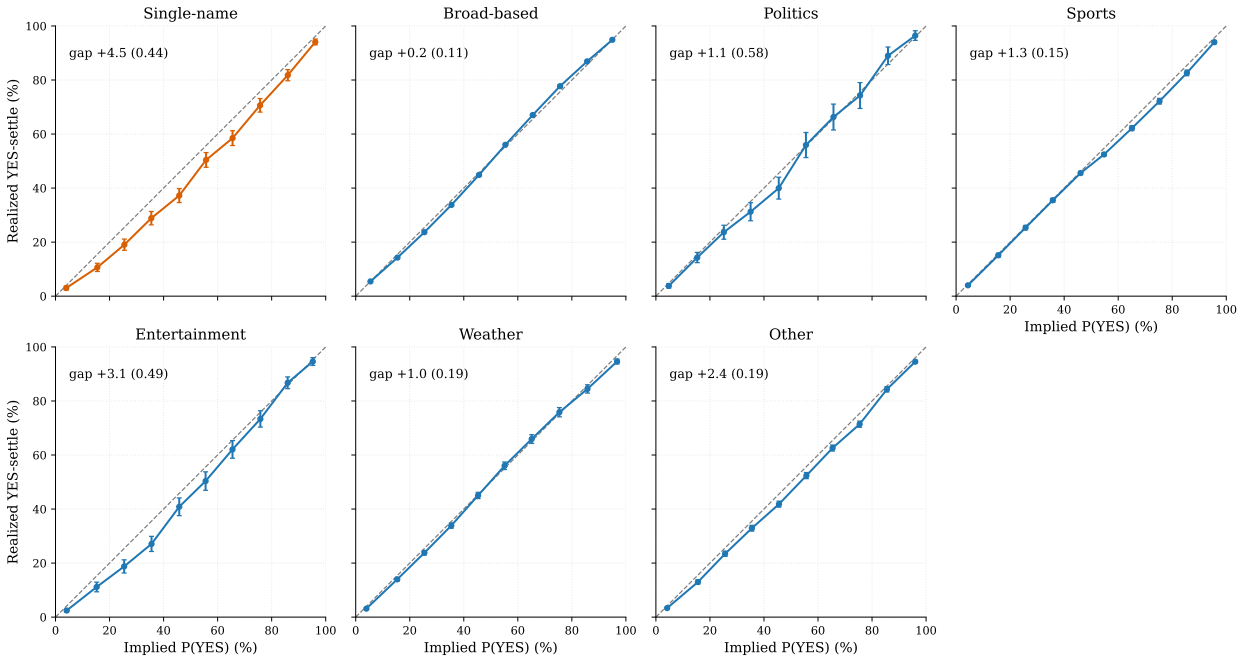

Figure B.1: Calibration by category. Each panel repeats the estimator of Figure 6 on one category of the paper's classification, with parlay markets excluded: markets with at least twenty trades in the first 80% of trading life contribute one unit of weight each, allocated across ten-cent price buckets in proportion to their own window volume. Vertical bars are 95% confidence intervals with standard errors clustered by market. Panel annotations report the equal-weighted per-market calibration gap (window price paid minus realized settlement, in percentage points) with its standard error.