Makers and Takers:
The Economics of the Kalshi Prediction Market

Constantin Bürgi* Wanying Deng† Karl Whelan‡

University College Dublin

January 2026

Abstract

Starting in 2021, Kalshi was the first federally licensed prediction market in the United States and remains the dominant platform in this segment. Using transaction-level data on over 300,000 contracts, we provide the first systematic evidence on its pricing. Contract prices are informative and become more accurate as markets approach closing, but they display a clear favorite–longshot bias: low-price contracts win far less often than required to break even after fees, while high-price contracts win more often and yield small positive returns. All trades are offered by Makers and accepted by Takers. Makers earn higher returns than Takers, but returns for both exhibit a favorite–longshot pattern. We interpret these facts using a framework that reflects Kalshi's quote-driven microstructure, in which agents sort into trading decisions based on their beliefs. We find that modest levels of disagreement and a small tendency to overstate small probabilities are sufficient to reproduce the favorite–longshot patterns in the data.

Keywords: Prediction Markets, Kalshi, Favorite–Longshot Bias, Market Microstructure, Bid–Ask Spreads

JEL Classification: G14, G23, G40

*constantin.burgi@ucd.ie

†wanying.deng@ucdconnect.ie

‡karl.whelan@ucd.ie. We would like to thank Oana Peia for comments and Prachi Srivastava and Jessica Whelan for help with obtaining the data used in the paper.

1. Introduction

Prediction markets match people who place wagers backing events to happen with people who wager that the event will not happen. Promoted in popular books by Surowiecki (2004) and Sunstein (2006), these markets have long been considered useful for their ability to summarize the public's information but their development was limited by regulatory restrictions in the United States. A 2008 article by 22 high-profile social scientists including four economics Nobel Prize winners (Arrow et al. 2008) recommended easing these restrictions and argued for "The Promise of Prediction Markets."

In November 2020, the Commodity Futures Trading Commission (CFTC) gave the startup firm Kalshi a license for a market of this type (technically, a designated contract market). Importantly, unlike previous prediction markets such as PredictIt or the Iowa Electronic Markets (IEM), Kalshi's license allows it to operate without stake limits. Kalshi (Arabic for "everything") operates markets across a huge range of topics such as culture, weather, company announcements, financial markets, politics and (since early 2025) sports. Notably, Kalshi called the 2024 US Presidential election correctly, predicting a high likelihood of a Trump win when most polling suggested the race was a toss-up.

Kalshi is a decentralized quote-driven market with price quotes ranging from 1c to 99c. Traders participate by either making offers to others to back Yes or No for an event or by accepting these offers. Kalshi describes the market structure as follows: *You can think of every trade on the platform as a deal taking place between two participants: a "maker" and a "taker". The maker is the first one to the table: They declare a side they're willing to buy (Yes or No), how much they're willing to pay, and how many contracts they are looking to buy at that price. Takers can see all available offers and match with the most generous one.*[1] A distinctive feature of Kalshi's data is that it records directly which side of each trade was initiated by the Maker and which by the Taker, so we can compare Makers and Takers without relying on noisy trade-initiator classification algorithms such as the Lee and Ready (1991) approach often used with stock market data.

This paper is the first to present systematic evidence on prices in the Kalshi market and also provides a theoretical model based on Kalshi's market microstructure to explain our findings. We present evidence on the outcomes of traded contracts on Kalshi since the market opened in 2021 through to April 2025. We have transaction-level data on 46,282 different Kalshi contracts on specific outcomes (e.g. Will Margot Robbie win the Oscar for best actress?) from 12,403 individual events (e.g. Who will win the Oscar for best actress?) each of which was open for at least 24 hours. We collect the final traded price as the market closed and also, where available, previous prices from 24-hour intervals up to 10 days before markets closed. This gives us over 300,000 prices when we factor in both sides of each traded contract.

We show that Kalshi's contract prices are relatively accurate predictors of outcomes and their

[1] https://help.kalshi.com/trading/who-are-you-trading-with

accuracy increases as markets come closer to closing. However, we also find that Kalshi's prices display a systematic favorite–longshot bias. Low-price contracts win less often than required to break even, with the opposite applying high-price contracts. In particular, investors who buy contracts costing less than 10c lose over 60 percent of their money. In contrast, there is statistically significant evidence that contracts with prices above 50c earn a small positive rate of return. These findings hold across a wide range of different categories of Kalshi's markets, though there is some evidence that the bias in prices is diminishing over time. Overall, the average rate of return on Kalshi contracts is about minus 20%.

We also use Kalshi's data on which side instigated trades to calculate returns on investment for both Makers and Takers. We find a favorite–longshot bias for contracts bought by both Makers and Takers but the pattern is more pronounced for prices accepted by Takers.

We present a model, adapted from Whelan (2025), to explain our findings. Previous theoretical models such as Gjerstad (2004) and Wolfers and Zitzewitz (2006) have assumed there is a single market-clearing price that equates supply and demand for contracts. In practice, quote-driven prediction markets feature two prices, one for accepting a Yes offer and one for accepting a No offer and there is typically a bid–ask spread separating these prices. We model the process of matching between Makers and Takers and illustrate how it endogenously produces the bid–ask spreads that you see on Kalshi. Previous work also largely ignored fees since early prediction markets such as Iowa Electronic Markets did not charge them. We model Kalshi's pre-2025 fee structure of charging a fee to Takers but not Makers.

Agents in the model have heterogeneous beliefs and self-select into trading decisions to be Makers or Takers based on those beliefs and the fee structure. Takers have more extreme beliefs than Makers and are willing to buy contracts at higher prices instead of being a Maker and risking not getting matched. This results in worse returns for Takers. To match the evidence of favorite–longshot bias for both types, we adapt the model to incorporate an over-estimation of low probabilities. A calibration with modest levels of disagreement among traders and a small over-estimation of low probabilities matches the findings in the data well.

The paper is organized as follows. Section 2 begins with a brief review of previous literature on prediction markets and describes how the Kalshi market operates. Section 3 presents our initial evidence on the performance of contracts and Section 4 provides our results on differences in outcomes for those who offer versus those who accept contracts. Section 5 presents our theoretical framework. Section 6 discusses why the apparent pricing anomalies we have found have not been competed away.

2. Previous Literature and How Kalshi Works

Here we provide a brief description of some of the relevant previous research on prediction markets and then discuss how the Kalshi prediction market works.

2.1. Previous Literature

The first prediction markets pre-dated the Internet. Forsythe et al. (1992) reported that the Iowa Presidential Stock Market, which operated on a computer network at the University of Iowa, yielded predictions of the vote shares in the 1988 US presidential election that outperformed opinion polls. Subsequent research on what became known as the Iowa Electronic Markets (IEM), such as Berg, Nelson and Rietz (2008) and Berg and Rietz (2019) further reported impressive findings on its predictive accuracy. Wolfers and Zitzewitz (2004) reported similarly strong predictive power for a number of other lower-profile prediction markets. To our knowledge, only one other academic paper has used data from Kalshi: Swanson, Wang and Wu (2025) analyze the impact of Fed announcements on prices on Kalshi's markets on macroeconomic data releases.

Of relevance for the results we present below, Berg and Rietz (2019) reported that IEM prices did not exhibit any favorite–longshot bias. Worth noting, though, is the evidence presented by Page and Clemen (2013) that prices from the InTrade prediction market tended to be too high for longshots and too low for favorites. However, this finding related to contracts that were traded well over 10 days in advance of the closing of the market. Page and Clemen described how the discounting of the potential future payout could explain this phenomenon and showed that there were no significant deviations between prices and win probabilities when contracts were close to closing. Our focus will be on contracts with no more than 10 days before closing. Of relevance, also, is Page's (2012) finding of the so-called Yogi Berra bias ("it ain't over till it's over") using InTrade, in which prices for losing teams in sports events in the final 15 minutes were too high relative to the fraction of times they won.

The theoretical literature on prediction markets has largely focused on whether it is reasonable to interpret a prediction market price as an unbiased estimate of the public's average belief about the probability of an event happening. Manski (2006) presented a simple model with risk-neutral investors who invested all their wealth in their preferred contract and showed that when observing a price p in a prediction market, the average belief could be as low as p^2 or as high as $2p - p^2$. Key contributions by Gjerstad (2004) and Wolfers and Zitzewitz (2006) showed that if agents set the size of their trading positions by maximizing log utility and the distribution of beliefs about a probability in the population were symmetric, then the prediction market price that equated demand and supply on either side of the contract was the average belief of participants. He and Treich (2017) showed that this result does not apply for any other CRRA utility functions.

2.2. How Kalshi Works

Kalshi operates a quote-driven market. Users post offers that are visible on the site and people can choose to either accept or reject these offers. This means there are always two potential prices at which a Yes contract can trade: one when a Maker offers other traders the opportunity to buy Yes and one when a Maker offers other traders the opportunity to buy No (in which case the Maker would be getting a Yes contract).

To illustrate how it works, consider a specific Kalshi event market: "What will monthly inflation be in the April 2025 US Consumer Price Index?" Figure 1 shows what the market looked like on May 7, 2025, six days prior to the CPI release, with the upper panel showing the historical traded prices. The figures quoted as percentages on the left-hand-side are the last traded Yes prices for each option, while the blue and purple boxes show the current best offers available to back Yes and No. For example, at a cost of 32c you could buy a contract that will pay out $1 if CPI inflation in April was over 0.3% and at a cost of 72c you could buy a contract that will pay out $1 if CPI inflation in April was 0.3% or under. Note that all trades occur at integer-valued numbers of cents, so it is not possible to buy a contract for 32.5c.

Figure 2 shows the options available when you click on the blue 32c button for Yes for Above 0.3%. It provides information on volumes in the order book for the "Above 0.3%" market. At this point in time, people had four different actions they could have taken in this market.

1. **Take the best available Yes offer of 32c**. The order book in this case shows you can buy $33.60 worth of Yes contracts at this price. If you bought more than this, you would have to pay 33c for the rest of your investment, with $3,336 available at that price. The 33c Yes price (and other more expensive offers to back Yes) shows up under "Asks" on the order book.

2. **Make an offer for a Yes contract for a price below 32c**. If you consider 32c too high, you can make an offer to buy a Yes contract at a lower price. For example, if you post a 30c offer for Yes, it becomes the best available price for someone seeking to back No (i.e. they would pay 70c for No) so 70c would show in the purple box in Figure 1. A key difference with this strategy is that accepting Yes at 32c means you definitely get your wager placed while seeking a cheaper price may mean your offer doesn't get matched.

3. **Take the best available No offer at 72c**. When you click on the purple 72c button, a second order book shows up. This has the same information as the Yes order book but the prices are quoted from the No perspective. We can see that $16.80 is available to buy No contracts at a price of 72c (because this is the amount made available by people seeking a Yes contract at 28c).

4. **Make an offer for a No contract for less than 72c**. For example, if you decided to seek a No contract at 69c, then you would have the best offer under "Yes" and it would be your 69c offer

that would show up as 31c in the blue box (because $1 - 69c = 31c). If that offer was accepted, you would have bought a No contract for 69c, which is cheaper than you could get from clicking on the No button and taking the 72c offer.

It is worth noting here that the last Yes contract price for Above 0.3% is 28c and this matches the most expensive bid price. This suggests the last trade likely involved someone accepting an offer from a Maker seeking a Yes contract for 28c. The transaction-level data provided by Kalshi explicitly identify which side of a trade was the Maker and which was the Taker.

Unlike early prediction markets like IEM, Kalshi charges fees. During the period corresponding to our sample, Kalshi charged a per-contract fee to people accepting offers (Takers) of $\$0.07P\,(1-P)$ where P is the price in dollars, rounding the total up to the nearest cent. They did not charge fees to Makers, i.e. people who posted limit orders that sat in the order book waiting to be matched. We incorporate fees for Takers into both our empirical return on investment calculations and our theoretical model. Kalshi began to charge fees on Makers after April 2025, which motivated our chosen cut-off point for sample.

Figure 1: A sample Kalshi market: CPI for April 2025

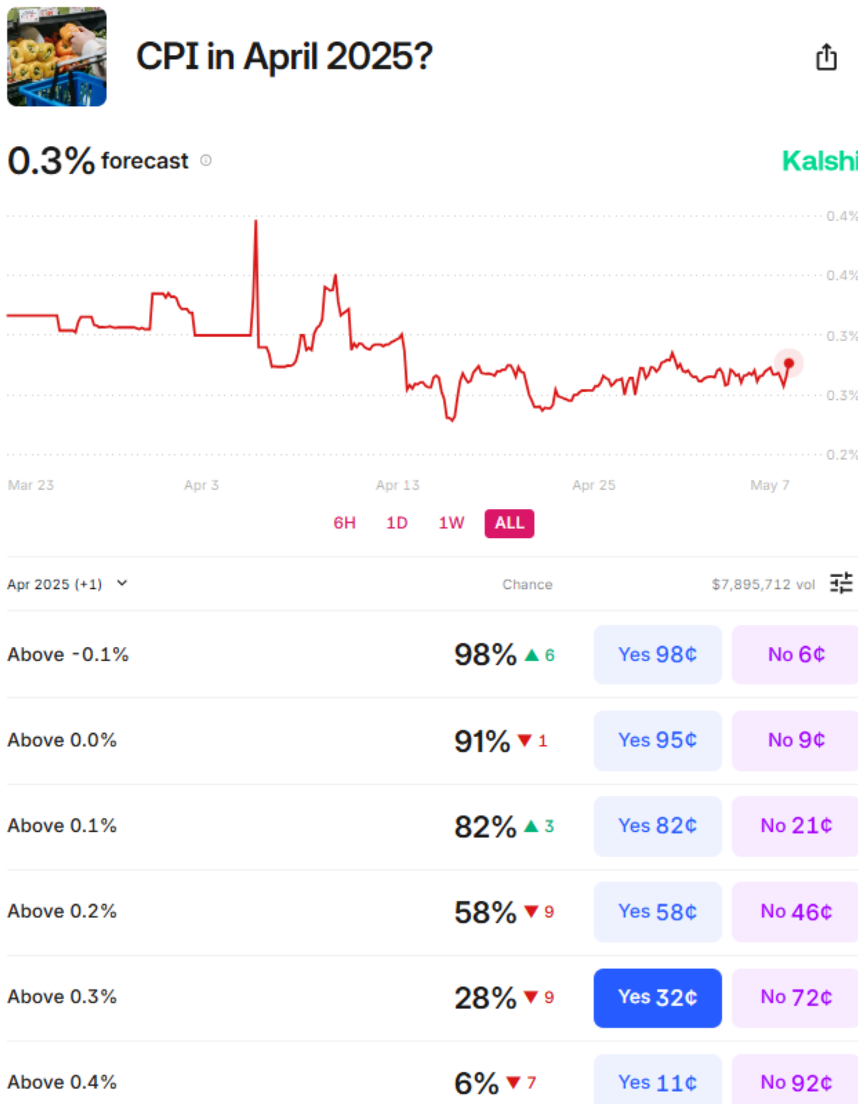


Figure 2: Options for betting for or against above 0.3%

	Price	Contracts	Total
	37¢	61	$6,859
	35¢	10,000	$6,836
	33¢	10,008	$3,336
Asks	32¢	105	$33.60
Yes	Last 28¢		
Bids	28¢	60	$16.80
	27¢	10,000	$2,717
	26¢	10,008	$5,319
	25¢	20	$5,324

3. Evidence on Kalshi's Prices

Here, we introduce our dataset and provide evidence on the predictive power of Kalshi's prices and returns on these investments. We start with summary statistics and then present regression evidence formally testing for a favorite–longshot bias.

3.1. Data and Summary Statistics

To obtain the data, we registered with Kalshi to get access to their Application Programming Interface (API). We used Python scripts to get transaction-level data from the API for contracts from Kalshi's inception in 2021 through April 2025. We focus only on contracts that have reached a total trading volume upon market closure of at least $1,000 to ensure there was a meaningful level of market activity. In addition, we drop contracts where the final bid–ask spread is larger than 20c.[2] Kalshi runs some markets that re-set every hour, generally markets that speculate on the value of cryptocurrencies or stock market indices at the end of the hour. We do not analyze these markets, limiting ourselves to markets that are open at least 24 hours.

The resulting dataset has the last traded Yes prices before the market closed for 12,403 events (e.g. Who will win the Oscar for best actress?) and 46,282 different Yes contracts (e.g. Will Margot Robbie win?) One aspect of the dataset worth noting is that volumes on Kalshi increased substantially in January 2025 as it began taking bets on sporting events. Our results below are not sensitive to cutting the data off in December 2024 or to excluding the category containing sports bets.

Going from final trades on the day a market closed back in 24-hour intervals to ten days before closing, we obtained a total of 156,986 prices for traded Yes contracts, so we have a maximum of 11 daily observations for each Yes contract and the average number of daily observations across Yes contracts is 3.5.[3] Associated with each Yes price is the corresponding No price that was taken by the other side, which means we have 313,972 prices for purchased contracts. We also obtained data on whether contracts were successful or not, which allows us to check the predictive accuracy of prices and the pre- and post-fee returns for all contracts.

As detailed in Table 1, the most common length of time a market is open in our dataset is 24 hours. Examples of these short-lived markets include Kalshi's markets for the top temperature in various cities tomorrow and markets like "Which app will be the most downloaded to iPhones tomorrow?" Other markets are available for longer. So, for example, we have 6,754 observations on Yes contracts that were traded every day back to 10 days before the market closed and 12,861 markets where we record the final traded price and the price from 24 hours earlier.

[2] Kalshi also reports final traded prices and volumes separately from their transaction-level data. We omitted 63 Yes contracts from the transaction-level data that did not match these figures.

[3] We collected these as follows. If, for example, the last trade on a market was at 5:57pm on the day the market closes, we also record the last trade before 5.57pm for each of the previous days that trades were available.

Some Kalshi events have simple Yes/No outcomes on a single event, so they have one Yes and one No: For example, Kalshi operates markets like "Will there be a government shutdown this year" and you can pick either Yes or No. Other markets have multiple mutually exclusive winners and you can back Yes or No for any of them: For example, "Who will be the most popular artist on Spotify this year?" Some, like the CPI market discussed earlier, offer multiple different Yes contracts available with their success dependent on a single published number or else markets like "Which phrases will be mentioned by the Fed chair during the FOMC press conference?" where there are many different phrases and no limit on how many can win. The average number of traded Yes contracts on the final day per event in our data was 3.7.

Table 2 shows the number of combined Yes and No contracts in our data for each of the 10 unit price intervals going from 1c to 10c up to 90c to 99c. It shows that our data contain far more contracts with extremely high or low prices than with mid-range prices. Of the 313,972 contract prices recorded, about two-thirds have prices either less than 10c or greater than 90c. There are only 8,351 contracts with prices between 50c and 59c.

Table 3 reports the average total final volumes traded by decile. Typical total volumes on Kalshi are low relative to most financial markets with the median amount of money staked on each contract (including money placed on both Yes and No) being $8,982. The distribution of volumes is highly skewed, with lots of markets having low volumes but those in the highest decile often having total volumes of over $1 million. Our data also report the size of individual transactions. The mean transaction size is $100 while the median is $35.

Table 1: Yes trade observations by number of days available before closing

Data Availability	Observations
Closing day only	19,338
Closing day and day before closing	12,861
Closing day and 2 days before closing	1,723
Closing day and 3 days before closing	1,083
Closing day and 4 days before closing	1,311
Closing day and 5 days before closing	1,288
Closing day and 6 days before closing	1,461
Closing day and 7 days before closing	159
Closing day and 8 days before closing	150
Closing day and 9 days before closing	154
Closing day and 10 days before closing	6,754
Total	46,282

Table 2: Total observations for all Yes and No contracts in each price range

Price Range	Number	Percent
1c-10c	106,209	33.8
11c-20c	20,395	6.5
21c-30c	12,558	4.0
31c-40c	10,049	3.2
41c-50c	7,199	2.3
50c-59c	8,351	2.7
60c-69c	10,049	3.2
70c-79c	12,558	4.0
80c-89c	20,395	6.5
90c-99c	106,209	33.8
Total	313,972	100

Table 3: Average final total volume deciles

Decile	Average Volume	Freq
1	$1,199	4,632
2	$1,735	4,634
3	$2,455	4,621
4	$3,593	4,626
5	$5,454	4,629
6	$8,481	4,628
7	$13,387	4,628
8	$21,286	4,628
9	$35,960	4,628
10	$526,245	4,628
Total	$61,977	46,282

3.2. Winning Fractions and Forecast Accuracy

As you can see from Figure 1, Kalshi's website reports its last traded Yes price in the form of percentages with the clear implication that, for example, if we examined a large sample of Yes contracts with prices equal to 60c, then we would expect the events being backed to happen 60% of the time. A simple way to assess this is to sort the data by price and record the fraction of wins for contracts at each price level.

Specifically, consider a contract with price P_{ij} in dollars which returns a payout of \$1 if event i ends in outcome j. Define

$$Y_{ij} = \begin{cases} 1 & \text{if outcome } j \text{ occurs} \\ 0 & \text{if outcome } j \text{ does not occur} \end{cases} \tag{1}$$

Figure 3 plots the empirical average win rate $\bar{Y}_{ij}$ for each possible value of P_{ij}, as well as a standard 95% confidence interval. We also show a 45 degree line. If this line lies inside the confidence intervals, then a standard t-test can't reject the hypothesis that the winning fraction was consistent with the prediction from the traded price.[4] One quick conclusion is that Kalshi's prices are pretty good predictors of outcomes. The winning fraction fluctuates around the 45 degree line, particularly in the middle price ranges where the sample sizes are smaller, but overall we can say that the higher a Kalshi price is, the more likely is the outcome.

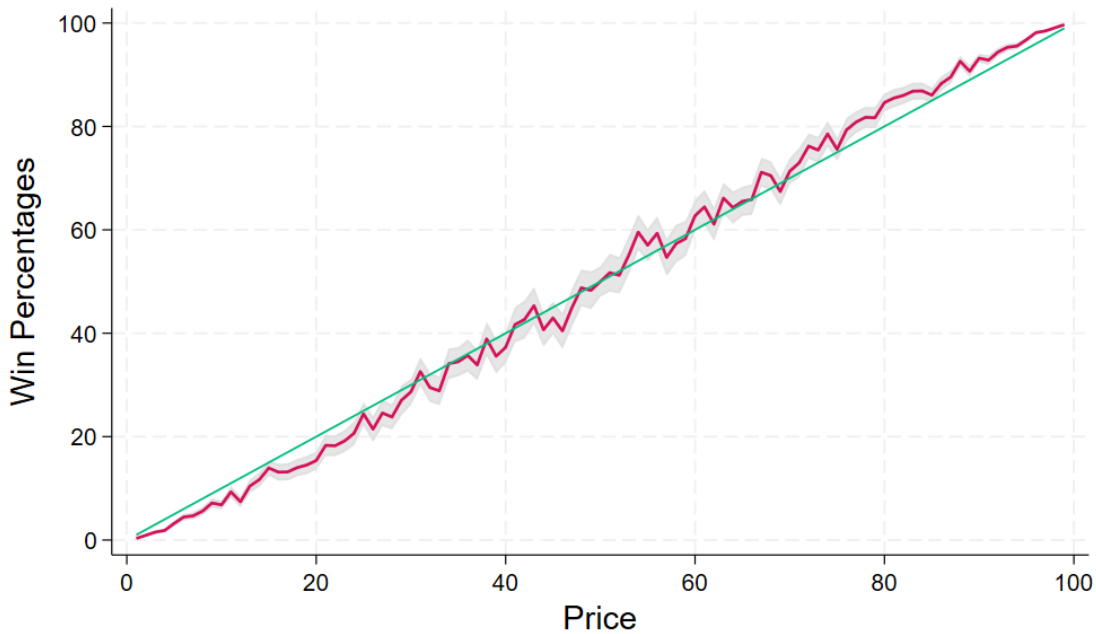
Interpreting Kalshi's prices as probabilistic forecasts, we can also assess the accuracy of these forecasts and whether they become more accurate as we get closer to the closing of the market. Evidence that Kalshi forecasts improve over time as more information becomes available is provided in Figure 4. We calculated the absolute error for each contract as $|Y_{ij} - P_{ij}|$ and then averaged them to obtain the mean absolute error (MAE) for various samples. The top-left chart in Figure 4 uses the data from all events where contracts were available from 10 days prior to market closing up to the closing price and reports the mean absolute error (MAE) of the forecast implied by the prices for each day. We can see that the MAE declines with each day as the market gets closer to closing. The decline is smooth until the last day, at which point there is a steep drop. This pattern is replicated in the other charts, which show MAE calculations for samples of contracts that were available for smaller numbers of days prior to the market closing.

While this is evidence of Kalshi's ability to forecast accurately, there is also a key shortcoming. Figure 3 shows a systematic pattern of contracts with relatively low prices having win rates (and typically also 95% confidence intervals) that lie below the 45 degree line, while the opposite applies for relatively high-price contracts. This mirrors the well-documented favorite–longshot bias in betting markets, which has not been recorded previously for relatively short-dated prediction markets such

[4] We have also estimated this relationship using bootstrapped kernel regression methods and found the same results.

as the one we are looking at.[5]

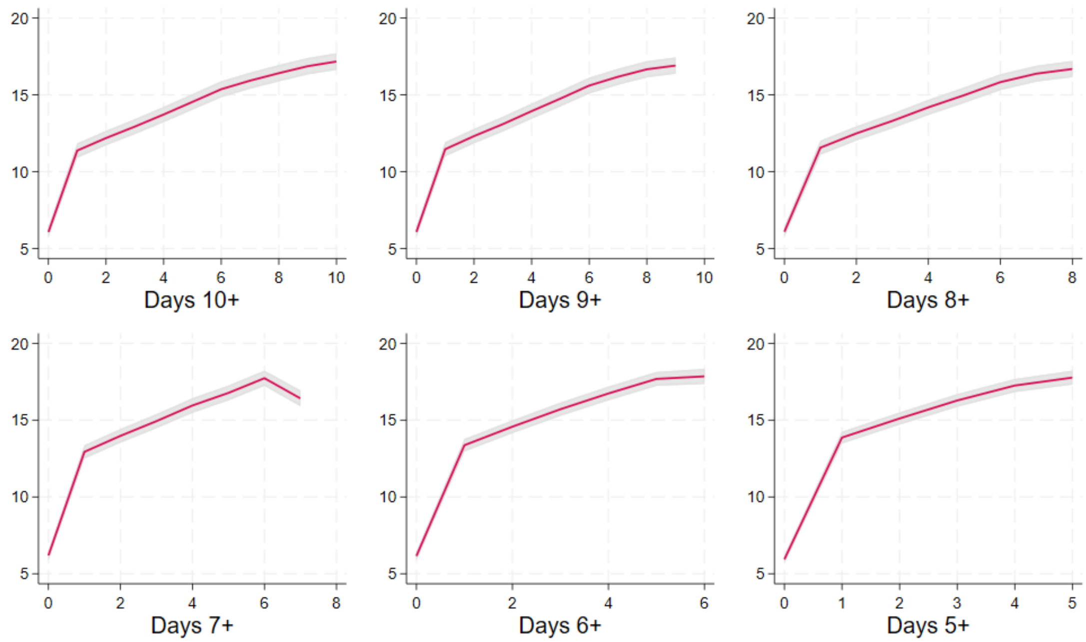
Figure 3: Win percentage sorted by price (shaded areas are 95% confidence intervals)



Notes: The figure shows the fraction of contracts that won for each price for the full sample of 313,972 Yes and No contracts.

[5]Snowberg and Wolfers (2008) and Ottaviani and Sørensen (2008) are excellent surveys of the theoretical and empirical literature on the favorite–longshot bias.

Figure 4: Mean Absolute Error by forecast horizon (shaded areas are 95% confidence intervals) for samples of contracts available up to 5 days out, 6 days out etc.



Notes: The figures show the MAE that won for each price for different sub-samples of the 313,972 Yes and No contracts for which prices were available continuously from the number of days prior to closing indicated.

3.3. Returns on Contracts

We also calculated ex post returns on contracts. Calculating the return for Makers is simple but there is a slight complication when calculating returns for Takers because of the fee they pay. We could calculate the rate of return on purchasing a single contract using the fee for a single contract. However, it is rare that anyone would buy only one contract. Kalshi calculates their fee as $\$0.07P\,(1-P)$ times the number of contracts where p is the price in dollars, rounding the total up to the nearest cent. This rounding means the average fee is generally slightly higher than Kalshi's stated fee rate for a single contract. We calculated an average fee for Takers based on the purchase of 100 contracts rounded up to the next cent. The round-up makes the fee on an average contract price of 50c equal to 1.77% of the price, slightly higher than the 1.75% it would be without the round-up.

Given our imputed commission rate for each contract, c_{ij}, we calculate the post-fee rate of return for Takers as

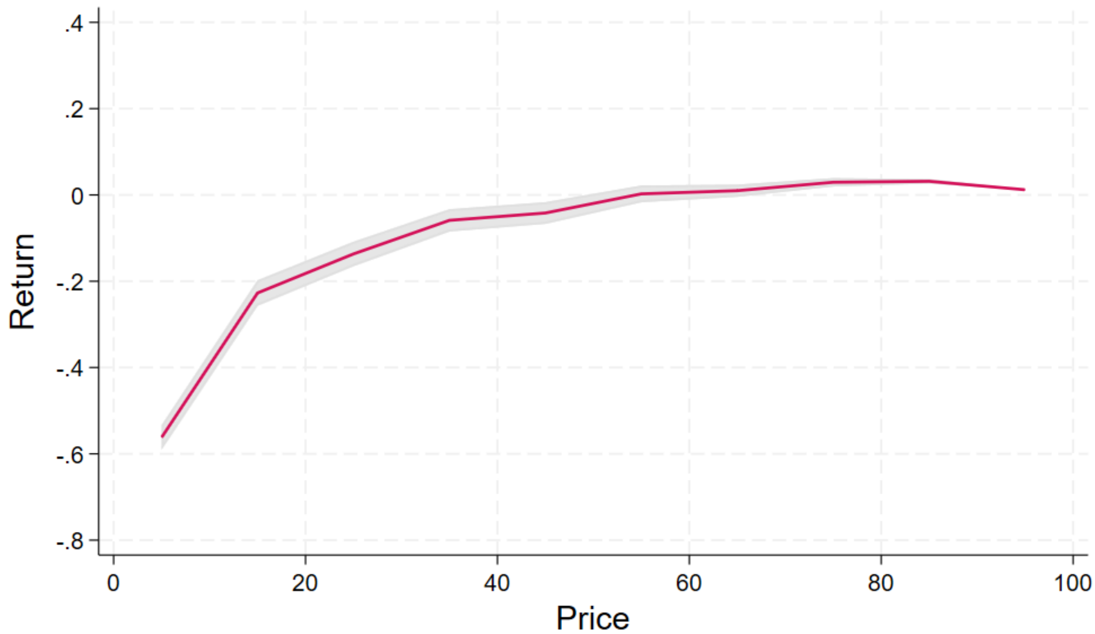
$$R^c_{ij} = \frac{Y_{ij} - P_{ij} - C_{ij}}{P_{ij} + C_{ij}} \tag{2}$$

In other words, when calculating the return once fees are considered, we treat the total investment to be the cost of the contract P_{ij} plus the associated commission.

Figure 5 shows the average returns on investment (incorporating fees for Takers) in contracts sorted by 10c-size price bands (1c-10c, 11c-19, ..., 91c-99c). The results are striking. The pattern of low-price contracts under-performing, which was evident in Figure 3, has a dramatic effect on their rate of return. Average loss rates for contracts costing 10c and under are over 60%. These loss rates fall as prices rise above 10c and there are small positive returns for contracts above 50c. For contracts above 70c, there is evidence of statistically significant, though small, positive post-fee returns.

These results are driven by the fact that, while the shortfall in win rates for low-price contracts shown in Figure 3 may appear small, it is a relatively large fraction of the price. A 5c contract that only wins 3% of time will have a minus 40% average return before fees. In contrast, the over-performance of higher-price contracts represents a relatively small amount of extra money relative to their prices. A 95c contract that wins 98% of the time has a pre-fee average return of 3.1%. Prior to fees, Kalshi participants simply swap money, so by definition the average return prior to fees across the total volume of money invested is zero. However, the asymmetric pattern of returns on cheap and expensive contracts means the average pre-fee return on a Kalshi contract in our data is -20%.

Figure 5: Rate of return on investment (incorporating fees for Takers) sorted by 10c-size price bands (shaded areas are 95% confidence intervals)



Notes: The figure shows the average rate of return on investment for various price-range subsamples of the 313,972 Yes and No contracts.

3.4. Regression Evidence

A formal way to test the accuracy of the forecasts implicit in Kalshi's traded prices is to use the classic Mincer-Zarnowitz (1969) regression

$$Y_{ij} - P_{ij} = \alpha + \psi P_{ij} + \epsilon_{ij} \tag{3}$$

The null hypothesis of the price being an unbiased predictor can be tested via the standard F-test of $\alpha = \psi = 0$. This approach has been used often to assess whether forecasts are unbiased, including the well-known contribution of Keane and Runkle (1990).[6] We report results from applying this approach to various sub-samples of our data in Tables 4 to 9.

A few technical aspects of these regressions are worth mentioning. First, because the variable being forecasted takes the value of either 0 or 1, this is effectively a linear probability model. These models can be objected to because they can predict probabilities outside the $[0, 1]$ interval but, in this case, the null hypothesis is that the expected value of Y_{ij} increases one for one with P_{ij} and not the

[6] A more recent application that relates to this example where probabilistic forecasts are assessed is Hegarty and Whelan (2024) which uses the Mincer-Zarnowitz regression approach to test for the accuracy of probabilities based on bookmakers' odds for soccer and tennis.

nonlinear relationships implied by Logit or Probit models. Also, any predicted value being outside $[0, 1]$ would itself be a violation of the null hypothesis.

Second, the regressions are restricted to the profits and prices of the 156,986 Yes contracts. For a Yes contract with price P_{ij} and profit $Y_{ij} - P_{ij}$, the price and profit for the No contract are just 100c minus the Yes contract values so the No observations are just a linear transformation of the Yes observations. This means a regression specification with both Yes and No contracts would double the apparent sample size without adding any useful information, thus distorting the standard errors downward.

Third, standard errors are clustered at the event-level and at the Yes-contract level, because there are negative correlations within observations on the same event (if one contract wins a mutually exclusive event, then the others did not) and because there are positive correlations between the observations across multiple days for the same Yes contract (if the Yes contract won, then the residual in the regression is positive for all the recorded observations on this contract from the first daily observation that we record up to the final price before the market closes.)

The first column in Table 4 reports the results from regression equation 3 for the full sample of 156,986 Yes contracts. As expected from the figures already provided, the rejection of the null hypothesis of prices as an unbiased predictor of outcomes is highly statistically significant. Indeed, the null hypothesis is firmly rejected for all sub-samples we present here. The form of the rejections fit with the favorite–longshot pattern observed in Figure 3. The intercepts are negative and the coefficient on the price is positive, indicating a pattern in which win rates are lower than predicted for low prices and then rise to being higher than predicted.

The remaining columns in Table 4 show the regression results for sub-samples for various different types of contracts. As noted above, some Kalshi markets have one contract for a specific event (e.g. is the temperature above X on a specific date and specific place?) Others have a series of related contracts for one event. The outcomes being considered can be mutually exclusive or instead have multiple possible winners as was the case with the CPI inflation market we showed earlier. For mutually exclusive contracts, we further distinguished between contracts where the description includes greater, less or in-between and classify these as numerical (e.g. GDP growth less than 2%, between 2-4% or greater than 4%). The remaining mutually exclusive contracts include events like who wins a particular Oscar. The results here show the hypothesis of Kalshi prices being unbiased predictors of outcomes is strongly rejected for all of the different types of contracts, with the ψ coefficient being largest for single contract markets.

Table 4: Regressions of pre-fee profit on price for different types of contracts

	(1)	(2)	(3)	(4)	(5)	(6)
			Non	Exclusive	Exclusive	All
	All	Single	Exclusive	Numerical	Other	Exclusive
Price	0.034***	0.045***	0.036***	0.014***	0.020**	0.017***
	(0.005)	(0.009)	(0.007)	(0.004)	(0.008)	(0.004)
Constant	-1.736***	-0.595	-1.839***	-1.637***	-1.916***	-1.756***
	(0.153)	(0.611)	(0.499)	(0.116)	(0.186)	(0.105)
Observations	156,986	16,433	58,602	46,674	35,277	81,951
F-test p-value	0.000	0.000	0.000	0.000	0.000	0.000

Notes: The table reports estimates of regression equation 3 in which the dependent variable is pre-fee profit of the specific contract. Single are single events with one contract, Non-Exclusive are non-mutually-exclusive multiple events (e.g. GDP larger than 2% and GDP larger than 3%). Exclusive Numerical are mutually exclusive events reliant on a numerical outcome and Exclusive Other are the remaining mutually exclusive events (e.g. who wins an Oscar). Standard errors clustered at event and contract level except for single contracts (which are only clustered at the contract level) in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5 shows separate regressions for the final traded Yes prices and the Yes prices drawn from the various one-day intervals prior to this, going back to 10 days before the market closed. While Figure 4 had shown prices becoming more accurate as the number of days until market closing declined, the regressions show the F-test still rejecting the null hypothesis for each of the days.[7]

Table 6 examines whether the amount of traded volume influences the favorite–longshot bias pattern that we have found. One could imagine that the favorite–longshot bias inefficiency might only be observed in markets with limited amounts of trading while arbitrageurs intervene in larger-volume markets to align prices better with winning rates. The table reports separate regressions for closing Yes prices for five quintiles based on the total final volume traded before market closing. The null hypothesis of prices being an unbiased predictor of outcomes is rejected for each of the quintiles, with a favorite–longshot bias pattern evident in the coefficients. The lowest volume quintile has the largest ψ coefficient but, other than that, there is no evidence of prices in higher-volume markets being more efficient predictors.

An alternative possibility could be that, rather than total market volume being what matters, prices could be more accurate when the size of the individual transactions is bigger because people pay more attention to fundamentals when they are placing more money at risk. Table 7 shows that this is not the case. Organizing the data into five quintiles by the average size of the transaction associated with each price, the null hypothesis is rejected for each quintile and, in fact, the quintile with the highest average transaction size has the largest ψ coefficient.

Table 8 reports regressions for separate categories of markets such as financials, crypto, climate and weather. Again, the null is rejected for all categories, though the ψ are smaller for politics and entertainment.

Finally, in Table 9, we report the regressions separately for each calendar year. Volumes on Kalshi have grown since 2021 and the market's users may have become more sophisticated. We find, however, that the null hypothesis of prices being an unbiased forecaster of outcomes is rejected for each of the years. There is some evidence of a weakening in the favorite–longshot bias because the ψ coefficient for the 2025 data is smaller and less statistically significant.

[7] Standard errors here were clustered only at the event level because each Yes contract only shows up once in these regressions.

Table 5: Regressions of pre-fee profit on price for different amounts of days prior to market closing

Days	Price	SE	Constant	SE	Observations	F-test p-value
0 Day	0.036***	(0.001)	-2.025***	(0.049)	46,282	0.000
1 Day	0.017***	(0.006)	-1.453***	(0.176)	26,835	0.000
2 Day	0.041***	(0.007)	-1.964***	(0.219)	13,996	0.000
3 Day	0.036***	(0.008)	-1.448***	(0.271)	12,281	0.000
4 Day	0.032***	(0.008)	-1.217***	(0.293)	11,201	0.000
5 Day	0.029***	(0.009)	-1.598***	(0.322)	9,918	0.000
6 Day	0.021**	(0.009)	-1.324***	(0.348)	8,631	0.000
7 Day	0.037***	(0.010)	-1.569***	(0.363)	7,178	0.000
8 Day	0.035***	(0.010)	-1.644***	(0.375)	7,024	0.000
9 Day	0.040***	(0.010)	-1.821***	(0.379)	6,886	0.000
10 Day	0.038***	(0.010)	-1.680***	(0.389)	6,754	0.000

Notes: The table reports estimates of regression equation 3 in which the dependent variable is the pre-fee profit of the specific contract. Standard errors clustered at event level in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6: Regressions of pre-fee profit on price for different final trading volume quintiles

	(1) Quintile 1	(2) Quintile 2	(3) Quintile 3	(4) Quintile 4	(5) Quintile 5
Price	0.045***	0.039***	0.032***	0.036***	0.032***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)
Constant	-2.422***	-2.242***	-1.903***	-1.870***	-1.668***
	(0.120)	(0.108)	(0.096)	(0.093)	(0.098)
Observations	9,266	9,247	9,257	9,256	9,256
F-test p-value	0.000	0.000	0.000	0.000	0.000

Notes: The table reports estimates of regression equation 3 in which the dependent variable is pre-fee profit of the specific contract for the last price of each contract. Quintile 1 are the 20 percent of contracts with the lowest volume, Quintile 5 the ones with the highest volume. Standard errors clustered at the event level in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Regressions of pre-fee profit on price for different mean transaction size quintiles

	(1) Quintile 1	(2) Quintile 2	(3) Quintile 3	(4) Quintile 4	(5) Quintile 5
Price	0.036***	0.033***	0.038***	0.034***	0.043***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Constant	-2.211***	-1.847***	-1.798***	-1.917***	-2.346***
	(0.130)	(0.104)	(0.105)	(0.097)	(0.089)
Observations	9,257	9,256	9,257	9,256	9,256
F-test p-value	0.000	0.000	0.000	0.000	0.000

Notes: The table reports estimates of regression equation 3 in which the dependent variable is pre-fee profit of the specific contract for the last price of each contract. Quintile 1 represents the 20 percent of contracts with the smallest mean transaction size, Quintile 5 the ones with the largest mean transaction size. Standard errors clustered at the event level in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8: Regressions of pre-fee profit on price by different categories of contracts

	(1) All	(2) Financials	(3) Climate & Weather	(4) Crypto
Price	0.034***	0.032***	0.031***	0.058***
	(0.005)	(0.004)	(0.005)	(0.014)
Constant	-1.736***	-1.431***	-0.997***	-1.944**
	(0.153)	(0.164)	(0.243)	(0.756)
Observations	156,986	27,123	29,924	8,150
F-test p-value	0.000	0.000	0.000	0.000

	(5) Politics	(6) Entertainment	(7) Economics	(8) Other
Price	0.022	0.020	0.034***	0.053***
	(0.021)	(0.012)	(0.010)	(0.009)
Constant	-1.912***	-2.809***	-0.978	-2.392***
	(0.340)	(0.364)	(0.972)	(0.467)
Observations	26,819	25,541	24,405	15,192
F-test p-value	0.000	0.000	0.000	0.000

Notes: The table reports estimates of regression equation 3 where the dependent variable is pre-fee profit of the specific contract. Columns split the sample by topical category (Financials, Climate & Weather, Crypto, Politics, Entertainment, Economics, Other). Standard errors clustered at the event and contract level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9: Regressions of pre-fee profit on price for different years

	(1)	(2)	(3)	(4)	(5)
	2021	2022	2023	2024	2025
Price	0.041***	0.023**	0.036***	0.048***	0.021*
	(0.015)	(0.011)	(0.009)	(0.006)	(0.011)
Constant	-1.649	-1.589***	-1.531***	-1.793***	-1.851***
	(1.224)	(0.366)	(0.357)	(0.289)	(0.253)
Observations	3,855	24,913	23,559	53,338	51,321
F-test p-value	0.026	0.000	0.000	0.000	0.000

Notes: The table reports estimates of regression equation 3 in which the dependent variable is profit of the specific contract. Standard errors clustered at the event and contract level in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

4. Evidence on Makers and Takers

Our data allow us to test the model's predictions about outcomes for Makers and Takers because, for every transaction, the data available from Kalshi's API records which side of a contract (either Yes or No) was the Taker's. This level of detail is unusual in empirical work on market microstructure. The standard Lee and Ready (1991) algorithm infers trade direction from quotes and transactions, but misclassifies 10–20% of trades in many applications (Odders-White 2000; Theissen 2001). Kalshi's direct identification of Makers and Takers therefore eliminates a major source of measurement error and provides a degree of precision that is rarely available in this literature.

Table 10 shows the number of observations for contracts bought by Makers for each price range as well as the total number of contracts. It shows that while Makers do buy many lower-priced contracts, they are more likely to buy higher-priced ones. This means the contracts that cost 10c or less, which perform particularly poorly, are mainly bought by Takers.

Table 10: Total observations for all Yes and No contracts in each price range and for contracts bought by Makers

Price Range	Total	Makers	Makers Share
1c-10c	106,209	46,185	43.5%
11c-20c	20,395	9,527	46.7%
21c-30c	12,558	6,136	48.9%
31c-40c	10,049	4,799	47.8%
41c-50c	7,199	3,565	49.5%
50c-59c	8,351	4,210	50.4%
60c-69c	10,049	5,250	52.2%
70c-79c	12,558	6,422	51.1%
80c-89c	20,395	10,868	53.3%
90c-99c	106,209	60,024	56.5%
Total	313,972	156,986	50.0%

Mincer-Zarnowitz F-tests, applied separately to Makers-only and Takers-only contracts, both firmly reject the null hypothesis of the traded price being an unbiased forecaster of outcomes. Again, the violation is due to low-cost contracts not winning often enough and higher-price contracts winning more often than needed to break even.

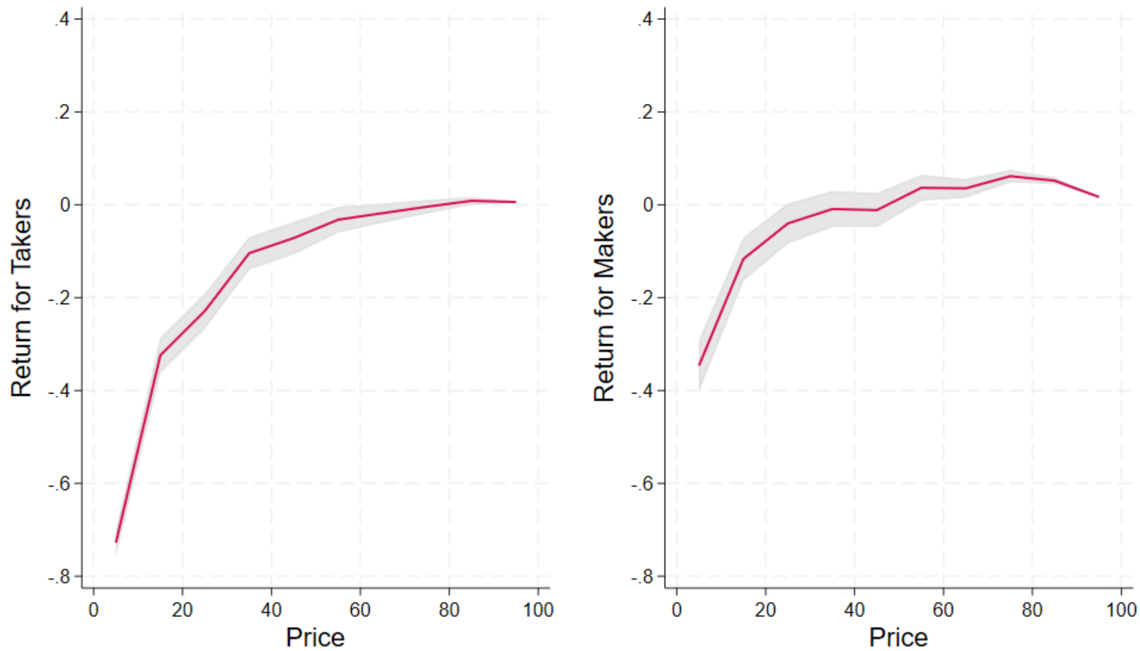
Figure 6 repeats the chart for average post-fee returns by price range, previously shown for all contracts in Figure 5, but this time separately with returns for Takers on the left and for Makers on the right. The first clear pattern is that average returns for Makers clearly exceed those for Takers: The average return on contracts for Makers was -9.64% while for Takers it was -31.46%. The F-statistic for the test of the null hypothesis that the average return for Makers equals the average return for Takers rejects the hypothesis with an extremely high level of significance. This finding is perhaps not surprising. By definition, Makers seek better prices than Takers but they must also be willing to cancel those orders if new evidence emerges that suggests they have made a bad offer.

Beyond average, Figure 6 show nonlinear patterns with increasingly negative returns as contract prices fall for both sides of trades but the pattern is more pronounced for Takers than for Makers. There is some evidence that Makers who buy contracts priced above 50c earn statistically significant positive returns but these returns are small relative to the huge loss rates for Takers who buy cheap contracts. On average, Makers who buy contracts costing 50c and over earn a 2.6% rate of return.

Because we include the same contract at different points during its lifetime, we want to make sure that our results are not driven by the small minority of contracts that are in the sample up to 11 times. At the same time, we want to assess whether this pattern identified in Figure 6 also holds during the various different days before closing. To this end, Figures 7 and 8 reproduce the charts for different times during the lifetimes of contracts. For example, the top-middle sub-figures show the returns using final outcomes with confidence intervals for all prices available 8 days before closing. For each of the days relative to closing, we see the nonlinear pattern of worsening returns for Takers as prices fall, with large and statistically significant loss rates for the prices of 10c and below and typically also for prices in the 11c-20c range.

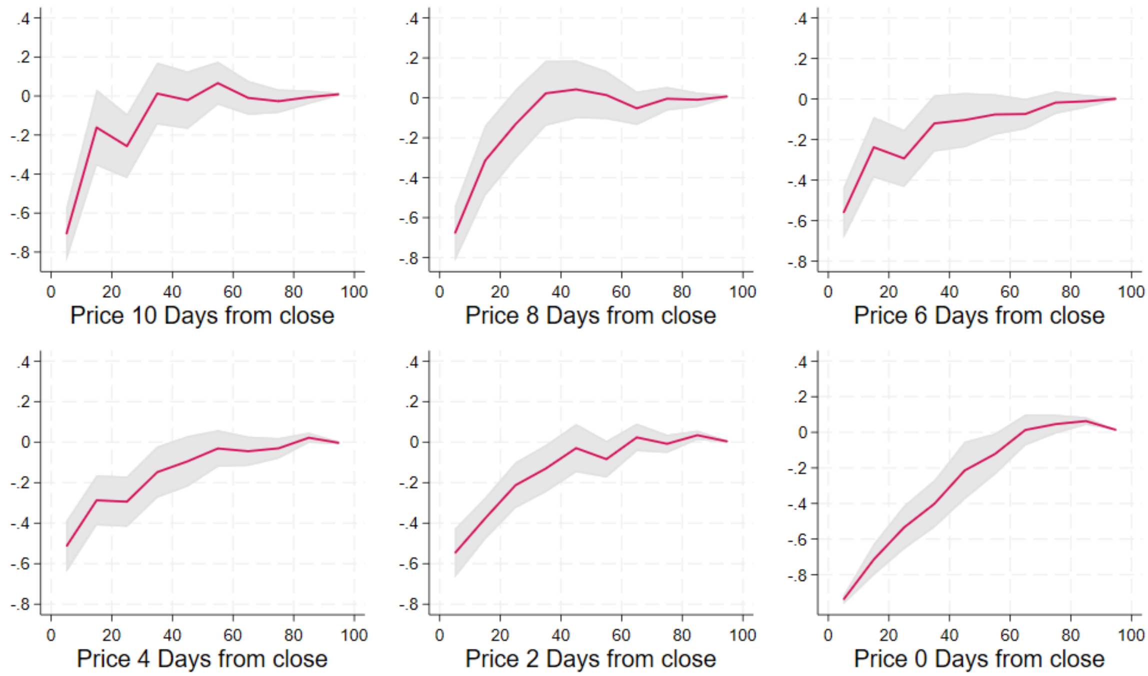
Returns for Makers also continue to show the pattern of returns generally worsening as the price of their contract falls, though most of these returns are not statistically significantly different from zero. The statistical significance of positive returns for Makers for higher-price contracts is generally not replicated here, most likely due to the smaller sample sizes and thus larger standard errors. Still, contracts that Makers buy for 10c or less have statistically significant negative rates of return for 5 of the 6 days shown here. Interestingly, the pattern of losses for Makers on closing prices (bottom left) is similar to those for Takers. This may point to over-optimism from Makers being a bigger issue as the market comes closer to closing than in previous days. Indeed, the larger losses for Makers and Takers on low-price contracts in the final day suggests a version of Page's (2012) "Yogi Berra effect" may be displayed in these markets.

Figure 6: Rates of return including commission by price (by 10c-size price bands) for Makers (right) and Takers (left) (shaded areas are 95% confidence intervals)



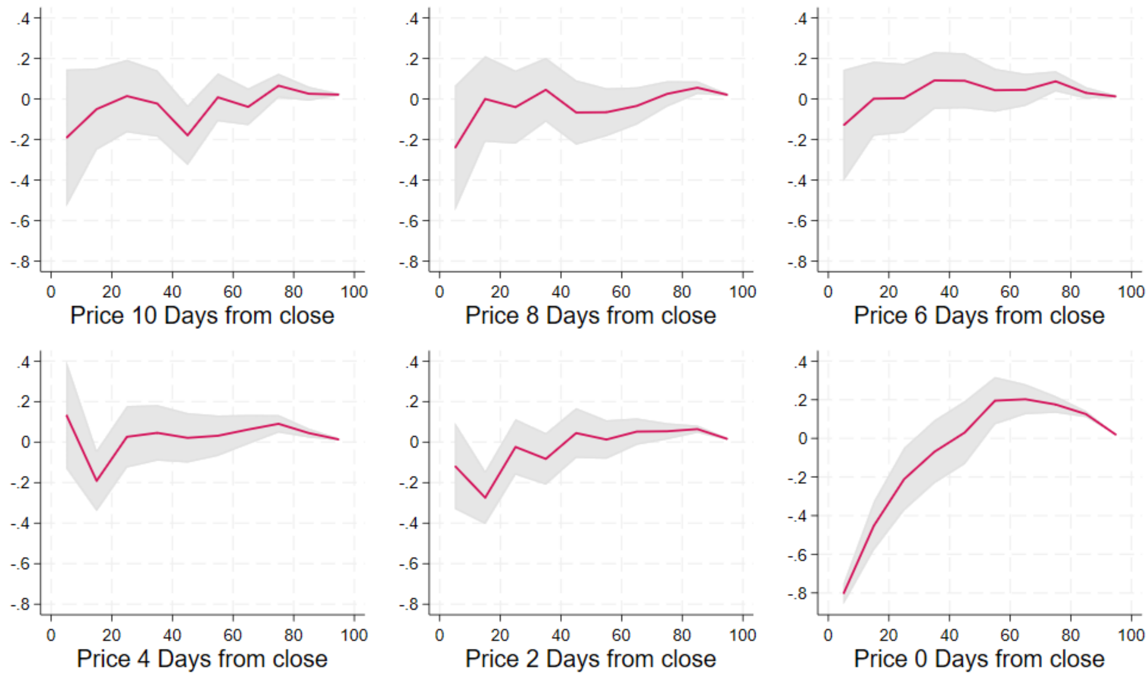
Notes: The left figure shows the average rate of return taking fees into account (as calculated by equation 2) for various price range sub-samples of the 156,986 contracts bought by Takers. The right figure shows the same for the 156,986 contracts bought by Makers.

Figure 7: Rates of return including commission by price (by 10c-size price bands) for Takers (shaded areas are 95% confidence intervals)



Notes: The figures show the average rate of return taking fees into account (as calculated by equation 2) for various price range sub-samples of the 156,986 contracts bought by Takers and for various numbers of days prior to the market closing.

Figure 8: Rates of return by price (by 10c-size price bands) for Makers (shaded areas are 95% confidence intervals)



Notes: The figures show the average rate of return for various price range sub-samples of the 156,986 contracts bought by Makers and for various numbers of days prior to the market closing.

5. A Model of Kalshi

To explain our findings, we adapt a model previously presented by Whelan (2025) and applied to the betting exchange markets that are popular in the UK. Betting exchanges quote decimal odds rather than prices between 1c and 99c and they have different fee structures, but the underlying trading structure is the same with Makers posting offers and Takers accepting them. We model the market as a simultaneous one-shot game, abstracting from intertemporal aspects. We have not included the option to post offers that would earn more for Makers than the current best offers available to Takers, nor have we modeled Kalshi's restriction of prices to integer values.

5.1. Assumptions

A Kalshi *Yes* contract pays $1 if the event occurs and zero otherwise. A *No* contract pays $1 if the event does not occur and zero otherwise. Participants differ in their beliefs about the probability of the Yes contract winning, with each person having their own subjective value of this probability π and the cumulative distribution function of beliefs being $F(\pi)$. The agents pick the trading option that they believe will give them the highest expected profit.

Kalshi quotes prices in cents, which we represent in dollars. Let

- $P_Y \in (0,1)$ denote the price a Taker pays to buy one *Yes* contract immediately;
- $P_N \in (0,1)$ denote the price a Taker pays to buy one *No* contract immediately.

Under the pre-2025 fee regime, only Takers pay fees. If a Taker executes at price P, the fee per contract is

$$f(P) = \gamma P(1-P), \tag{4}$$

where $\gamma > 0$ is a cost parameter, equal to 0.07 during the period we have examined. Makers pay no fee.

A trader may alternatively act as a Maker and post an order which is matched with some probability. We assume:

- A Maker who posts to buy *Yes* at price P_Y^M is matched with probability θ_Y.
- A Maker who posts to buy *No* at price P_N^M is matched with probability θ_N.

We assume traders buy a single standard-sized contract, so supply and demand for contracts are proportional to the fraction of agents willing to buy them. This assumption of an exogenous stake size is common in theoretical models of sports betting (see, for example, Shin 1991, Ottaviani and Sørensen 2010 and Hegarty and Whelan 2026). In contrast, the literature on prediction markets, such as Gjerstad (2004) and Wolfers and Zitzewitz (2006), has more commonly used an endogenous stake size, generally set in accordance with the Kelly criterion implied by log utility. We don't use this approach here because the evidence above shows that the typical investment on Kalshi is modest and very unlikely to be consistent with Kelly staking.

Matching is assumed to be randomly assigned, so for both Yes and No contracts, the probability of a Maker being matched with a Taker is just the ratio of Takers to Makers. The model's equilibrium is a set of decision rules, prices and matching probabilities that are consistent with all agents maximizing their subjective expected profit.

5.2. Decision Rules and Match Rates

Here we describe how traders make different decisions based on their beliefs.

Buying Yes: take immediately or post as a Maker

Consider an agent with belief π about the probability that *Yes* occurs. If they buy *Yes* immediately as a Taker at price P_Y, they pay the quadratic fee $f(P_Y)$ and their expected profit is

$$\pi - P_Y - f(P_Y) \tag{5}$$

Alternatively, they can post an order to buy *Yes* as a Maker. A posted order is executed with probability θ_Y. When it executes, it trades against a Taker who buys *No* at price P_N, which is equivalent to buying *Yes* at price $1 - P_N$. The expected profit from posting to buy *Yes* is therefore

$$\theta_Y \left(\pi - (1 - P_N)\right) \tag{6}$$

Posting to buy *Yes* is profitable (conditional on execution) only if $\pi \geq 1 - P_N$.

Buying *Yes* immediately is preferred to posting as a Maker whenever

$$\pi - P_Y - f(P_Y) \ \geq \ \theta_Y \left(\pi - (1 - P_N)\right) \tag{7}$$

This condition is satisfied if

$$\pi \ \geq \ \pi_Y^T \equiv \frac{P_Y + f(P_Y) - \theta_Y(1 - P_N)}{1 - \theta_Y} \tag{8}$$

Thus, agents with sufficiently strong beliefs buy *Yes* immediately as Takers, while those with $1 - P_N < \pi < \pi_Y^T$ prefer to post and accept the execution risk.

Buying No: take immediately or post as a Maker

If the agent buys *No* immediately as a Taker at price P_N, they pay the fee $f(P_N)$ and their expected profit is

$$(1-\pi) - P_N - f(P_N) \tag{9}$$

Alternatively, they can post an order to buy *No* as a Maker. A posted order is executed with probability θ_N. When it executes, it trades against a Taker who buys *Yes* at price P_Y, which is equivalent to buying *No* at price $1 - P_Y$. The expected profit from posting to buy *No* is therefore

$$\theta_N\left((1-\pi) - (1-P_Y)\right) = \theta_N\left(P_Y - \pi\right) \tag{10}$$

Posting to buy *No* is profitable (conditional on execution) only if $\pi \leq P_Y$.

Buying *No* immediately is preferred to posting as a Maker whenever

$$(1-\pi) - P_N - f(P_N) \;\geq\; \theta_N\left(P_Y - \pi\right) \tag{11}$$

This is satisfied if

$$\pi \;\leq\; \pi_N^T \equiv 1 - \frac{P_N + f(P_N) - \theta_N(1-P_Y)}{1-\theta_N} \tag{12}$$

Thus, sufficiently pessimistic agents buy *No* immediately as Takers, while those with $\pi_N^T < \pi < P_Y$ prefer to post.

Matching rates

Recalling that the CDF of beliefs is $F(\pi)$, the mass of agents who post to buy *No* as Makers is

$$F(P_Y) - F(\pi_N^T), \tag{13}$$

while the mass who buy *Yes* immediately as Takers is $1 - F(\pi_Y^T)$. Random matching implies

$$\theta_N = \frac{1 - F(\pi_Y^T)}{\left[F(P_Y) - F(\pi_N^T)\right]} \tag{14}$$

Similarly, the mass who post to buy *Yes* as Makers is

$$F(\pi_Y^T) - F(1-P_N) \tag{15}$$

while the mass who buy *No* immediately as Takers is $F(\pi_N^T)$. Random matching implies

$$\theta_Y = \frac{F(\pi_N^T)}{\left[F(\pi_Y^T) - F(1-P_N)\right]} \tag{16}$$

5.3. Multiple Equilibria and Sorting

The model can be summarized by four equations: two equations defining the cut-off thresholds for Taking versus Making (equations 8 and 12) and two equations defining the matching rates (equations 14 and 16). However, the model has six endogenous variables: the two quoted prices (P_Y, P_N), the two threshold beliefs $\left(\pi_N^T, \pi_Y^T\right)$ and the two matching rates (θ_Y, θ_N). With four equations and six unknowns, the system is underdetermined and admits multiple equilibria.

In particular, the model contains both thick and thin market equilibria. One way to frame this is in terms of the bid–ask spread, the gap between the prices faced by Takers and the posted prices sought by Makers. When matching probabilities are high, Makers are more confident that their posted orders will be executed, so a smaller price improvement is required to justify posting rather than trading immediately. This results in smaller gaps between Taker prices (P_Y, P_N) and Maker prices $(1 - P_N, 1 - P_Y)$ and thus narrower bid–ask spreads. By contrast, when matching probabilities are low, the lower likelihood of execution must be offset by more favorable posted prices for Makers, resulting in wider effective spreads.

The coexistence of thick and thin equilibria is a common feature in models involving search and matching, with Diamond (1982) providing a foundational example. Similar forms of multiplicity arise in contexts such as urban economics (Gautier and Teulings, 2009) and marriage markets (Burdett and Coles, 1997). A common feature that our model shares with these other examples is that the attractiveness of the decision to post an order depends on how many others are likely to accept it.

In what follows, we treat the matching probabilities θ_Y and θ_N as exogenous, setting both equal to an identical value θ, and choose values for these and the other parameters that match the observed returns in different price regions for Makers and Takers. Once we specify these parameters and the distribution function F, we have four nonlinear equations in four unknowns and solve the model using MATLAB's nonlinear least squares solving function `lsqnonlin`. The resulting equilibrium in the model is defined as a set of posted prices and thresholds such that individual decisions defined by the thresholds are consistent with both the prices and the exogenously-specified match rates. For a given set of sensible parameters, the model solves quickly and produces clear outcomes in which agents sort into one of five distinct groups as π goes from low to high: (i) Taking No at P_N (ii) Making by posting No at $1 - P_Y$ (iii) Making on both sides (iv) Making by posting Yes at $1 - P_N$ (v) Taking Yes at P_Y.

5.4. Modeling Beliefs

The final piece of the model is the specification of the distribution of beliefs about event probabilities. We assume that agents differ in their subjective assessments. For a true Yes probability of π^*, beliefs are distributed as $\pi \sim N(\mu(\pi^*), \sigma^2)$ where, motivated by the standard linear shrinkage approach to belief formation (Gabaix, 2014, Coibion and Gorodnichenko, 2015), the public's average belief is a convex combination of the truth π^* and an anchor of 0.5 as follows

$$\mu(\pi^*) = \pi^* + \beta(0.5 - \pi^*) \tag{17}$$

Here, $\beta \in [0,1]$ is a bias parameter. When $\beta = 0$, beliefs are unbiased on average ($\mu = \pi^*$). When $\beta > 0$, the mean belief is shifted toward 0.5, meaning agents systematically overestimate small probabilities and underestimate large probabilities. This specification captures a well-documented pattern in probability judgment (Kahneman and Tversky, 1979), where people tend to overweight unlikely events. As we discuss below, a model with only belief heterogeneity ($\beta = 0$) cannot predict the results reported in Figure 6 in which returns are systematically more negative for cheap contracts for both Makers and Takers.

This specification also implies consistent beliefs about an event and its complement since $\mu(\pi^*) + \mu(1 - \pi^*) = 1$. For example, if $\beta = 0.1$ and π^* = 0.15, the mean of the probability distribution for beliefs is μ = 0.185 while the corresponding mean belief for the event not happening is 0.815. This is important since the agents in our model will be considering both Yes and No contracts for every value of π^*.

The assumption that the standard deviation of beliefs is the same value σ, whatever the true probability π^*, is a highly simplified one. It allows some people to have subjective probabilities that are either negative or greater than one. In practice, for our preferred calibrations, this has little influence since these beliefs are uncommon and the belief thresholds for high and low values of π^* are all between zero and one. This means the few agents in the model with "impossible" beliefs just take the same decision to be Takers as others that have extreme but feasible beliefs. The greater issue is whether the inflexibility of this specification prevents the model from fitting the data well and we will show now that it does not.

5.5. Model Calibration and Fit

Our approach to estimating the model's fit and choosing illustrative parameter values is to match a very specific set of moments: the data on average returns on investment for Makers and Takers across different price ranges shown in Figure 6.

Specifically, to measure model fit, we compute a weighted distance between the model's predicted expected returns (averaged across the different underlying values of π^*) and the observed returns across the ten price bins for both Makers and Takers. Let m^{data} denote the vector of observed average returns (ten elements for Takers followed by ten elements for Makers) and $m^{model}(\theta, \sigma, \beta)$ denote the corresponding model predictions. Our measure of fit is

$$f(\theta, \sigma, \beta) = (m^{data} - m^{model})' W (m^{data} - m^{model}) \tag{18}$$

where W is a diagonal weighting matrix. Each bin is weighted by its share of total trading volume, so that bins with more trades receive greater weight in the measure of fit. The extreme price bins (1–10 cents and 90–99 cents) together account for 68% of trades and thus receive the highest weight.

One approach would be to use simulated method of moments estimation to report the best fit. In practice, however, we found that a wide range of parameter combinations produce nearly equivalent fits to the data, which is a common issue with this kind of highly nonlinear model.[8] Rather than report potentially non-robust "best" estimates, we performed a grid search across many parameter values and noted the patterns in the values that provided the best-fitting predictions.

The results showed that changes in the matching rate parameter θ and the belief dispersion parameter σ had similar impacts on model fit. For a fixed σ, higher match rates increase the ratio of Takers to Makers, which is achieved through narrower spreads, so that Taking becomes more attractive. Conversely, for a fixed θ, higher belief dispersion naturally generates more traders with extreme views who prefer to be Takers, but maintaining a fixed ratio of Takers to Makers requires wider spreads to discourage this additional taking. These parameters thus have offsetting effects on equilibrium spreads. For this reason, among the parameter combinations that have the top five percent of best fits, θ and σ are consequently highly correlated (0.80).

In contrast, the simulations suggest the probability over-weighting parameter β is tightly identified. Among the same set of best-fitting combinations, the values of β ranged only from 0.06 to 0.12. This tight range indicates that some degree of probability misperception is necessary for the model to match the observed favorite–longshot bias patterns.

Given this, we present a calibration that uses $\beta = 0.09$, which lies in the middle of the well-fitting parameter range. We set the matching rate as $\theta = 0.60$, consistent with Kalshi being a moderately liquid market where Makers have a reasonable execution probability. With these parameters fixed, we calculated the best-fitting disagreement parameter as $\sigma = 0.107$.

Figure 9 shows that this calibration fits the data well for both Makers and Takers across the full range of price deciles. In the model, the over-optimism about small probabilities leads to large losses for cheap contracts for both Makers and Takers. But losses are larger for Takers (note the different ranges for the two charts) for two reasons: Makers get better prices than Takers and Takers pay fees. These two elements affect Takers of very cheap contracts more. Kalshi's fee formula has a higher cost per price paid when the price is low. And the fixed disagreement width assumption in the model means bid–ask spreads are larger as a fraction of price for cheap contracts, thus driving a greater wedge between Maker and Taker returns.

As prices increase, the model matches the data in predicting that loss rates for Takers decline but they still lose money on average. It also matches the predictions of small profits for Makers on

[8] Weak parameter identification with flat regions of the objective function is a common feature of nonlinear structural models. See Canova and Sala (2009) for discussion.

contracts with higher prices. Makers benefit from the Takers having more extreme beliefs, from the underlying bias toward small probabilities in the distribution of beliefs and from not having to pay fees.

None of the calibrations with $\beta = 0$ fit well but to illustrate the comparative statics of the role this parameter plays, Figure 10 shows the results when we keep the other two parameters from our preferred calibration but set the bias in average beliefs to zero. Without biased beliefs, the model predicts that Makers should always be able to take advantage of the more extreme beliefs of Takers to earn profits. And, in the reverse of what we actually see in the data, it predicts that those returns should be biggest for cheap contracts. The assumption of a fixed disagreement variance drives this result. Takers with very high values of π are willing to pay a few cents above π^* for their contracts. That is a small fraction of their overall investment, so Takers buying expensive contracts don't lose that much, but this error represents a big fraction of the Maker's cheap price, so Maker expected returns are highest for low-price contracts and then fall as the price rises.

We can see from these results that the systematic bias in average probability beliefs is crucial to explaining the patterns seen in Kalshi prices but the overall level of this bias is quite modest. A model in which traders maximize their subjective expected profits with modest levels of disagreement can combine with a relatively weak pattern of biased beliefs to generate some very bad returns for cheap contracts, particularly for Takers.

Figure 9: Average expected return rates for Makers and Takers for $\beta = 0.09$, $\theta = 0.6$, $\sigma = 0.107$, sorted by 10 price bins and compared with the data from Figure 6

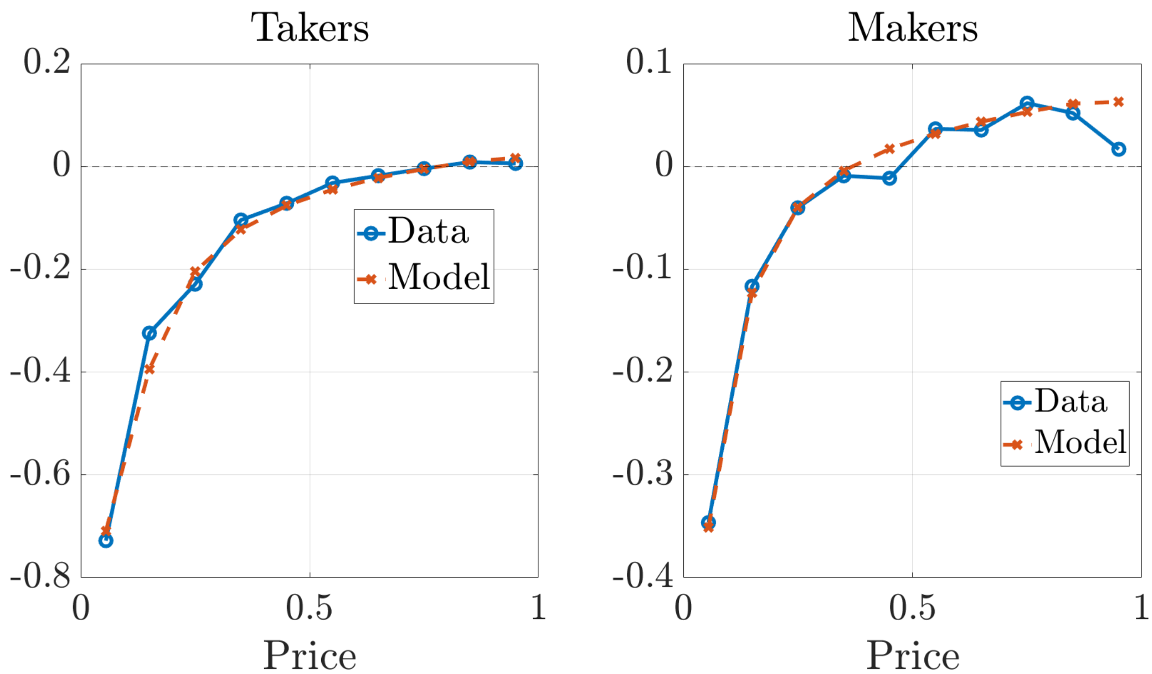
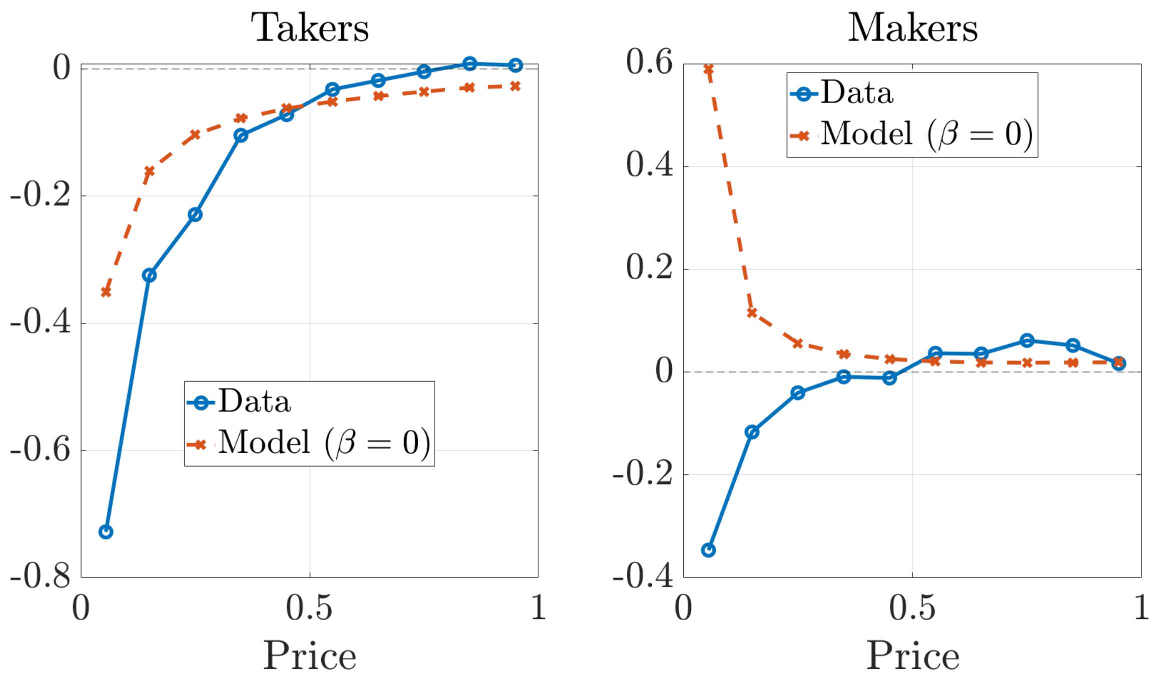


Figure 10: Average expected return rates for Makers and Takers for $\beta = 0$ (no bias), $\theta = 0.6$, $\sigma = 0.107$, sorted by 10 price bins and compared with the data from Figure 6



6. Why Aren't Biases Competed Away?

An interesting question about our results is why investors have not come in to undercut the profits made on higher-priced contracts by Makers and thus eliminate the favorite–longshot bias. We have noted that the rate of return after commission for Makers on contracts costing 50c and over is 2.6%. This may seem modest when compared with annualized returns on most financial instruments but, of course, it is not an annualized return. Someone could buy a Kalshi contract a few days before closing, make profits and then re-invest those profits. So there could be potential for large cumulative returns. We can point to three potential reasons why this hasn't happened up to now.

Small Volumes

The first potential reason is the small volumes on Kalshi's markets. Even its largest-volume markets are small relative to the kinds of markets that professional investors will typically be willing to participate in. Table 3 showed average final trading volume in the top decile of Kalshi markets was only $526,245. And at any point in time, the amount of liquidity available is far smaller: Notice the relatively small amounts available in the order book for the April CPI shown in Figure 2. This likely means that someone who wanted to invest substantial capital as a Maker seeking to buy high-price contracts may have to post prices that are less advantageous to Makers than the typical trades that we recorded here. This would further reduce the potential size of investments that could be made to exploit the bias in prices documented here. Furthermore, some of the attempts to trade as a Maker would not be matched, again reducing the overall amount actually invested.

Riskiness

The second potential reason is the riskiness of investments in Kalshi contracts. The asymmetric payoff structure of these contracts means that the standard deviations of returns dwarf the averages. The standard deviation of the rate of return on contracts bought by Makers costing 50c and over is 33%. This huge variance may deter many investors from buying these contracts.

A counter-argument is that a strategy of making lots of these investments could rely on the Law of Large Numbers to essentially guarantee profits with relatively little risk. However, Paul Samuelson (1963) famously argued that this idea ran counter to expected utility theory: If you rejected one gamble, you should also reject two independent versions of the same gamble. Pratt and Zeckhauser (1987) called this property of rejecting multiple independent gambles that would be rejected on a standalone basis "proper risk aversion" and demonstrated that this property applied to all power utility and logarithmic utility functions, i.e. those used in the vast majority of modern financial economic theory. It is thus possible that rational investors who have examined Kalshi's market have concluded that the positive expected returns for Makers on higher-price contracts represent no more than an appropriate return for the risk taken.

Lack of Information

The third potential reason is that perhaps participants in Kalshi's markets have not been aware of the favorite–longshot bias pattern that we have documented or the evidence for positive returns for Makers when buying higher-priced contracts. We think it will be interesting to see if the biases and return patterns that we have reported persist now that they have been publicly documented or whether they follow the pattern of other financial market anomalies in tending to disappear once they have been publicized (see Zaremba, Umutlu and Maydybura, 2020 and Shanaev and Ghimire, 2021).

7. Conclusion

This is the first academic paper to present a systematic study of prices on the Kalshi prediction market. Kalshi represents an important development in the history of prediction markets. Previous evidence had shown prediction markets could produce highly accurate prices but historical markets were limited in scope and had very small trading volumes. Kalshi runs a very wide range of markets and has uncapped trading volume, features proponents argued would improve predictive performance.

Our results show that, as of yet, this promise has not been fulfilled. Kalshi's prices exhibit a strong favorite–longshot bias, with systematic underpricing of high-probability events and overpricing of low-probability events. Returns are substantially negative for cheap contracts and modestly positive for more expensive ones. This pattern persists despite uncapped trading volumes and mirrors the biases long documented in traditional betting markets where prices are set by bookmakers or via a pari–mutuel system.

The data on positions taken by Makers and Takers allow us to go further than previous studies and examine how these pricing distortions are distributed across different types of traders. We find that both Makers and Takers earn negative returns on cheap contracts, but Takers lose substantially more. This differential pattern provides insight into the mechanism generating the favorite–longshot bias. Our model shows that the combination of belief disagreement, modest probability overweighting, and Kalshi's fee structure can explain both the overall bias and the Maker–Taker differential. Crucially, we demonstrate that behavioral bias is necessary—a model with heterogeneous but unbiased beliefs cannot match the observed patterns. But the level of bias required for our model to explain the data is modest. The large negative returns on some contracts do not necessarily imply that Kalshi has lots of traders with highly biased opinions.

These findings have implications for the use of prediction markets in policy and business contexts. While they are clearly a useful tool for aggregating information, our results suggest their prices should not be interpreted as unbiased probability estimates. The systematic biases we document are predictable and could in principle be corrected, but users should be aware that market prices—particularly for low-probability events—reflect both information and behavioral distortions. Despite hopes that deeper and wider markets would produce better prices, it may be that favorite–longshot bias is an inherent feature of markets where traders with heterogeneous beliefs and probability judgments interact with each other.

References

[1] Arrow, Kenneth J. et al. (2008). "The Promise of Prediction Markets," *Science*, Vol 320, Issue 5878, pages 877–878.

[2] Berg, Joyce, Forrest Nelson and Thomas Rietz (2008). "Prediction Market Accuracy in the Long Run," *International Journal of Forecasting*, Volume 24, pages 285–300.

[3] Berg, Joyce and Thomas Rietz (2019). "Longshots, Overconfidence and Efficiency on the Iowa Electronic Market," *International Journal of Forecasting*, Volume 35, pages 271–287.

[4] Burdett, Ken and Melvyn G. Coles (1997). "Marriage and Class," *Quarterly Journal of Economics*, Volume 112, pages 141–168.

[5] Canova, Fabio and Luca Sala (2009). "Back to Square One: Identification Issues in DSGE Models," *Journal of Monetary Economics*, Volume 56, pages 431–449.

[6] Coibion, Olivier and Yuriy Gorodnichenko (2015). "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts,"*American Economic Review*, Volume 105, 2644–2678.

[7] Diamond, Peter (1982). "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, Volume 90, pages 881–894.

[8] Forsythe, Robert, Forrest Nelson, George Neumann and Jack Wright (1992). "Anatomy of an Experimental Political Stock Market," *American Economic Review*, Volume 82, pages 1143–1161.

[9] Gabaix, Xavier (2014). "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, Volume 129, pages 1661–1710.

[10] Gautier, Pieter and Coen Teulings (2009). "Search and the City," *Regional Science and Urban Economics*, Volume 39, pages 251–265.

[11] Gjerstad, Steven (2004). Risk Aversion, Beliefs and Prediction Market Equilibrium. University of Arizona Working Paper 04-17.

[12] He, Xue-Zhong and Nicolas Treich (2017). "Prediction Market Prices under Risk Aversion and Heterogeneous Beliefs," *Journal of Mathematical Economics*, Volume 70, pages 105–114.

[13] Hegarty, Tadgh and Karl Whelan (2024). "Comparing Two Methods for Testing the Efficiency of Sports Betting Markets," *Sports Economics Review*, Volume 8, paper 100042.

[14] Hegarty, Tadgh and Karl Whelan (2026). "Market Structure and Prices in Online Betting Markets: Theory and Evidence," *Oxford Economic Papers*, Volume 78, pages 90–113.

[15] Kahneman, Daniel and Amos Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, Volume 47, pages 263–291.

[16] Keane, Michael P. and David E. Runkle (1990). "Testing the Rationality of Price Forecasts: New Evidence from Panel Data," *American Economic Review*, Volume 80, pages 714–735.

[17] Lee, Charles and Mark Ready (1991). "Inferring Trade Direction from Intraday Data," *Journal of Finance*, Volume 46, pages 733–746.

[18] Manski, Charles (2006). "Interpreting the Predictions of Prediction Markets," *Economic Letters*, Volume 91, pages 425–429.

[19] Mincer, Jacob and Arnold Zarnowitz (1969). "The Evaluation of Economic Forecasts" in *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance.* New York: National Bureau of Economic Research.

[20] Odders-White, Elizabeth R. (2000). "On the Occurrence and Consequences of Inaccurate Trade Classification," *Journal of Financial Markets*, Volume 3, pages 259–286.

[21] Ottaviani, Marco and Peter Norman Sørensen (2008). "The Favorite–Longshot Bias: An Overview of the Main Explanations," in *Handbook of Sports and Lottery Markets*, Elsevier.

[22] Ottaviani, Marco and Peter Norman Sørensen (2010). "Noise, Information, and the Favorite–Longshot Bias in Parimutuel Predictions," *American Economic Journal: Microeconomics*, Volume 2, pages 58–85.

[23] Page, Lionel (2012). "'It Ain't Over Till it's Over' Yogi Berra Bias on Prediction Markets," *Applied Economics*, Volume 44, pages 81–92.

[24] Page, Lionel and Robert Clemen (2013). "Do Prediction Markets Produce Well-Calibrated Probability Forecasts?," *Economic Journal*, Volume 123, pages 491–513.

[25] Pratt, John and Richard Zeckhauser (1987). "Proper Risk Aversion," *Econometrica*, Volume 55, pages 143–154.

[26] Samuelson, Paul (1963). “Risk and Uncertainty: A Fallacy of Large Numbers,” *Scientia*, Volume 98, pages 108–113.

[27] Shanaev, Savva and Binam Ghimire (2021). “Efficient Scholars: Academic Attention and the Disappearance of Anomalies,” *The European Journal of Finance*, Volume 27, pages 278–304.

[28] Shin, Hyun Song (1991). “Optimal Betting Odds Against Insider Traders,” *Economic Journal*, Volume 101, pages 1179–1185.

[29] Snowberg, Erik and Justin Wolfers (2008). “Examining Explanations of a Market Anomaly: Preferences or Perceptions?” in *Handbook of Sports and Lottery Markets*, Elsevier.

[30] Sunstein, Cass (2006). *Infotopia: How Many Minds Produce Knowledge*. Oxford University Press.

[31] Surowiecki, James (2004) *The Wisdom of Crowds*. Abacus.

[32] Swanson, Eric, Renxuan Wang and Yanbin Wu (2025). The Effects of Monetary Policy on Macroeconomic Expectations: High-Frequency Evidence from Traded Event Contracts, working paper, China Europe International Business School, Shanghai.

[33] Theissen, Erik (2001). “A Test of the Accuracy of the Lee–Ready Trade Classification Algorithm,” *Journal of International Financial Markets, Institutions and Money*, Volume 11, pages 147–165.

[34] Whelan, Karl (2025). Agreeing to Disagree: The Economics of Betting Exchanges, CEPR Discussion Paper Number 20633.

[35] Wolfers, Justin and Eric Zitzewitz (2004). “Prediction Markets,” *Journal of Economic Perspectives*, Volume 18, pages 107–126.

[36] Wolfers, Justin and Eric Zitzewitz (2006). Interpreting Prediction Market Prices as Probabilities. NBER working paper 12200.

[37] Zaremba, Adam, Mehmet Umutlu and Alina Maydybura (2020). “Where Have the Profits Gone? Market Efficiency and the Disappearing Equity Anomalies in Country and Industry Returns,” *Journal of Banking and Finance*, Volume 121. Article 105966.