Bias in Prediction Markets

by Liam Deaton

Abstract

This paper models the microeconomic foundations of "winner-take-all" prediction markets in the presence of outside investment options. I propose a distribution for belief heterogeneity, and model probability calibration under conditions where participants are risk neutral and risk averse. Market participants have heterogenous prior beliefs about the probability of event occurance. Trades get made when a buyer can be matched with a seller. Market participants have reliable outside options. I find that in equilibrium, prices may deviate substantially from the true probability of event occurance.

Introduction

Prediction markets are rapidly growing in public discourse around events and future predictions. Political pundits, equity researchers and news organizations all have begun to cite the odds presented in such markets. Many organizations have begun used prediction markets internally to improve forecasting. However, there is relatively little theoretical research around these markets, their efficiency, and deeper functional aspects of their potential for bias, especially as these markets reach maturation.

There are a variety of possible formats for prediction markets, but emergent popularity has centered around a "winner takes all" market, in which a contract on a binary Yes or No event is bought at a price confined to between \$0 and \$1 dollar. The winner of the bet receives \$1 if the bet resolves in their favor, otherwise recieves zero. Recent additions to the prediction market landscape that follow this pattern include Kalshi, Robinhood, PredictIt, and Polymarket, the latter trading in US dollar-backed crypto currencies. Collectively these markets trade billions. The prices of events, such as the winner of the 2024 election, are often understood to represent the market believed probability of an event occuring. The widespread reporting of these prices may also influence beliefs and behavior in the wider economy, though the extent to which this is the case is unknown.

The companies HP, and Siemens, and Google have all experimented with internal prediction markets to drive management. Understanding underlying market theory is vital for management to harness and rely on such internal markets.

There is limited evidence on the efficiency of winner-takes-all markets. Berg, Forsyth, Nelson and Reitz (2001) find that the Iowa Electronic Markets, a small-scale index election market where trades are limited to \$500, yield highly accurate predictions, and outperformed large-scale polling organizations. Evidence from a few other small cap markets favor this. However the motivation of agents in smaller markets may not be reflective of larger cap mature markets that we see today.

On the theory side of things, Justin Wolfers and Eric Zitzewitz (2005), build a model, testing how the bias between prices and true probability occurs for different utility functions under a variety of distributions. My paper improves upon their modeling by considering the effect of outside options, and time until event expiration. Additionally, I alter distributional assumptions, imposing a beta function with participant liquidity centered around mean beliefs.

Model

Participants are able to buy "yes" and "no" bets for an event A. A "yes" bet will yield a payout of 1 if the event occurs and a payout of 0 otherwise. For any given event, price of a "yes" bet is p and price of a "no" bet is $(1-p)$. Yes and no bets sum to one.

For a transaction to occur, when a "yes" bet is placed it must be matched to a "no" bet, otherwise the transaction does not occur.

Market participants have prior beliefs $\theta \in (0,1)$ that represent their perceived probability of the event occuring. The prevalance of beliefs are represented by a function $f(\theta) \in (0,1)$.

For a given market participant, their expected return of a "yes" bet is $Return(A_{yes}) = \theta$ and expected return of a "no" bet is $Return(A_{no}) = (1-\theta)$. The percentage return of a yes bet is thus $\frac{\theta}{p}$ and for a no bet is $\frac{(1-\theta)}{(1-p)}$

In addition to being able to place a bet on the event. Market participants are able to use their money to participate in an outside option such as bonds, which reliably can multiply their money at $(1+r)^t$ where t is the time until maturity of both the outside option and the prediction event, and r is the interest rate.

Belief distribution

For my experiments I will model belief distributions as a beta distribution due to it's flexible nature. I will assume that the distribution of beliefs have a mean that is equal to the "True" probability of the event, μ. In addition, I will set a standard deviation for the distribution s, representing the spread of the belief distribution. Thus our beta distribution has shape parameter defined by $B(\mu \times \frac{\mu(1-\mu)}{s^2}, (1-\mu) \times \frac{\mu(1-\mu)}{s^2})$.

Risk Neutral

Under true risk neutraility, market participants are trying to maximise their expected return as a function of price p and will bet if and only if the expected percentage return exceeds the outside option $(1+r)^t$.

Thus risk neutral market participants have 3 options, "yes", "no" and "outside option".Thus participants bet yes if $\frac{\theta}{p} > (1+r)^t$ no if $\frac{(1-\theta)}{(1-p)} > (1+r)^t$ and does not place a bet otherwise.

Under this model, the minimum type θ that is indifferent to taking a yes bet occurs at $\theta = p(r+1)^t$ and the maximum type that is indifferent to a "no" bet is $\theta = p(r+1)^t + (r+1)^t - 1$.

Under this model the number of yes bets taken can be modelled as

$$\begin{cases} 0 & \text{for} : p\left(r+1\right)^{t} \geq 1 \\ \int\limits_{p\left(r+1\right)^{t}}^{1} f(\theta) d\theta & \text{otherwise} \end{cases}$$

and the number of no bets taken is:

$$\begin{cases} 0 & \text{for} : p\left(r+1\right)^{t} - \left(r+1\right)^{t} \leq -1 \\ \int\limits_{0}^{p\left(r+1\right)^{t} - \left(r+1\right)^{t} + 1} f(\theta)\, d\theta & \text{otherwise} \end{cases}$$

This leads to the market equilibrium price occuring when they are equal, at.

$$\begin{cases} 0 & \text{for} : p\left(r+1\right)^{t} \geq 1 \\ \int\limits_{p\left(r+1\right)^{t}}^{1} f(\theta), d\theta & \text{otherwise} \end{cases} = \begin{cases} 0 & \text{for} : p\left(r+1\right)^{t} - \left(r+1\right)^{t} \leq -1 \\ \int\limits_{0}^{p\left(r+1\right)^{t} - \left(r+1\right)^{t} + 1} f(\theta)\, d\theta & \text{otherwise} \end{cases}$$

Substituting $f(\theta)$ with our beta function yields the equation:

$$0 = \begin{cases} 0 & \text{for}\ p\left(r+1\right)^{t} \geq 1 \\ \int\limits_{p\left(r+1\right)^{t}}^{1} \frac{\theta^{\frac{\mu^2(1-\mu)}{s^2}-1}(1-\theta)^{\frac{\mu(1-\mu)^2}{s^2}-1}\left(\frac{\mu^2(1-\mu)}{s^2}+\frac{\mu(1-\mu)^2}{s^2}-1\right)!}{\left(\frac{\mu(1-\mu)^2}{s^2}-1\right)!\left(\frac{\mu^2(1-\mu)}{s^2}-1\right)!}\, d\theta & \text{otherwise} \end{cases} - \begin{cases} 0 \\ \int\limits_{0}^{p\left(r+1\right)^{t} - \left(r+1\right)^{t} + 1} \frac{\theta^{\frac{\mu^2(1-\mu)}{s^2}-1}(1-\theta)^{\frac{\mu(1-}{s}}}{\left(\frac{\mu(1-\mu)^2}{s^2}\right.} \end{cases}$$

I am interested in how true probability μ differs from the equilibrium prices under this model. These pricing conditions are not solvable analytically, thus I computationally derive the curve for market clearing conditions below. Figure 1 shows the equilibrium prices across true probabilities μ, holding s constant. Notice how as $(1+r)^t$ becomes larger, the prices tend towards having an equilibrium price closer to 0.5 than is the true probability. This is due to increased difficulty finding buyers at more extreme probabilities. Additionally, increasing s makes the line straighter.

Risk Averse

In model 2 I add an additional assumption, which is that market participants are risk averse. While in the first scenario it was assumed that every participant, regardless of type staked their entire endowment into the market (or that each person places a single bet), I now assume that participants are risk averse, and are trying to maximize the expectation of their log returns. The mathematics of this problem is well-trodden, and the maximization assumption is solved by participants following kelly's formula, with discounted future returns. Under these conditions investors with a lower expected return will stake a smaller share of their endowment on the problem.

If an agent believes that the market has overpriced odds: **Share to bet** $= -\frac{p\left(1-\theta\right)\left(r+1\right)^{t}}{\theta} + \theta$

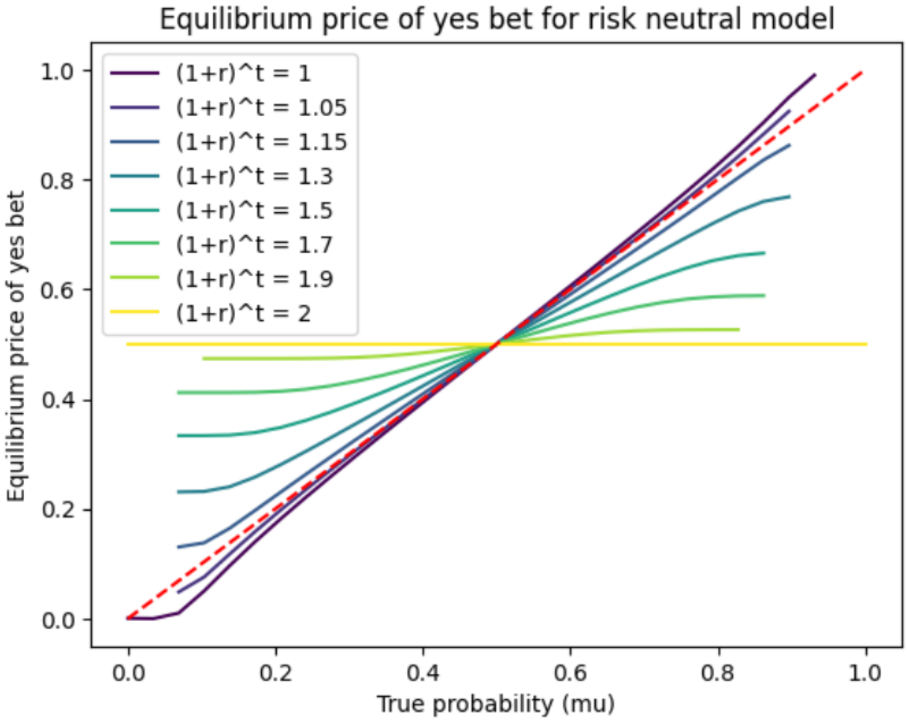

Figure 1: Equilibrium Price Of Yes in Risk Neutral Model

otherwise they take the opposite position, also using kelly's formula

Thus the market clearing conditions are:

$$\int_0^1 \begin{cases} \left(-\frac{p(1-\theta)(r+1)^t}{\theta}+\theta\right) f(\theta) & \text{for } \frac{p(1-\theta)(r+1)^t}{\theta}-\theta \leq 0 \\ \left(\frac{\theta(1-p)(r+1)^t}{1-\theta}+\theta-1\right) f(\theta) & \text{for } \frac{\theta(1-p)(r+1)^t}{1-\theta}+\theta < 1 \\ 0 & \text{otherwise} \end{cases} d\theta$$

Computationally solving for p for all possible $\mu \in (0,1)$ after subbing in our beta specification yields the curve in Figure 2.

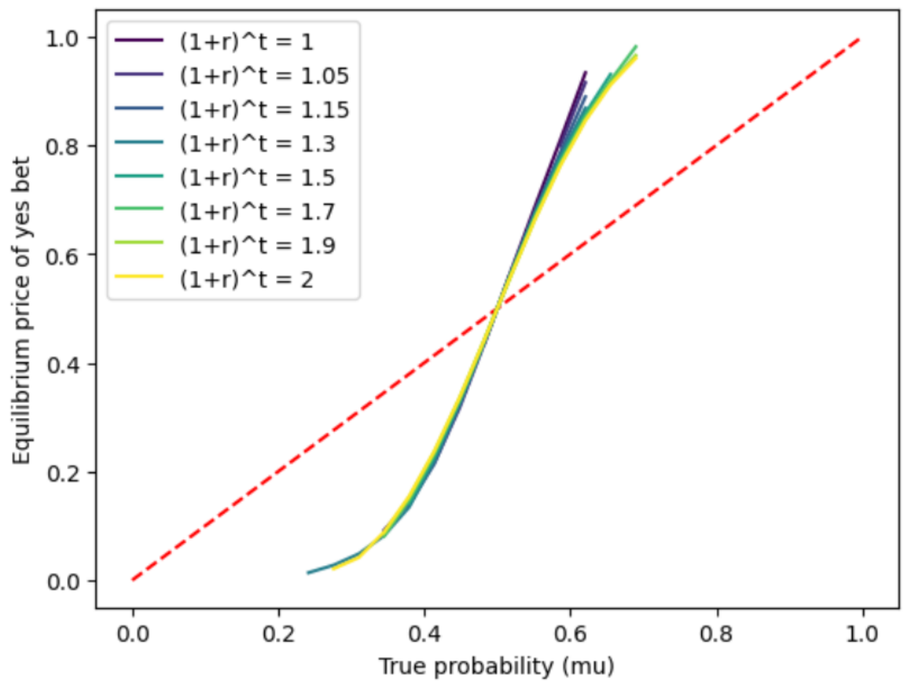

Figure 2: Price Curve for Risk Averse Model

Notice that as $(1+r)^t$ increases the ends begin to twist, however the curve changes less dramatically.

Conclusion

This paper provides a microfounded model to asses the bias in prices versus the true probability of an event occuring in "Winner-takes-all" prediction markets. Adding important twists including outside options to investor decision making.

This research is important. In the 2024 election many people from political pundits, to equity researchers, to news orginizations, have referenced the odds given on various prediction markets to better understand the true probability of election events. As these markets entrench themselves in mainstream discourse it is vital to understand how outside options and the interest rate bias these markets.

I believe Model 2 is likely the better charachterization of the most recently popularized, high cap markets. It is more consistent with the unusual distribution seen in polymarket order books, in which there are few orders around the price, and then a large increase in unfilled orders at a specific cutoff point on both sides of the price, after which unfilled orders rise in both directions.

Some important caveats exist however. There is much research on the psychology of gamblers that suggest that a reasonably large share of market actors may be less than risk averse, or even risk loving. Future research should consider markets that contain heterogenous risk preferences accross belief types, and how this affects equilibrium prices.

Additionally, this research potentially provides insight into how prediction markets are affected by the currencies used. There are existing prediction markets that are denominated in crypto-currencies, internal point systems, and in real dollars. The denomination of currencies significantly affect the interest rate of the outside option. For internal point systems, there are no outside options other than participating in prediction markets, so $(1+r)^t$ is exactly 1. Crypto-currencies may see extreme fluctuation in $(1+r)^t$ depending on the currency, trend in its particular price, ability to transfer funds into other currencies, and the slate of reliable investment instruments that exist in this space. The outside options in real dollars are relatively well known (t-bills). This can help explain why odds vary wildly across prediction markets denominated in different currencies, and why arbitrage seemingly does not occur.

Such considerations should also be used by prediction market designers, when they consider their use case. For example, internal company prediction markets under a “winner takes all” system might be more accurate if they are denominated in some sort of internal points system, with restrictions on payouts.