Are Prediction Markets Bayesian?

Matthew vonAllmen

Pitzer College

Are Prediction Markets Bayesian?

Introduction

Probabilities, in a Bayesian framework, refer to subjective beliefs about the state of the world. Bayesian agents update these beliefs in response to new information according to specific rules. If prediction market prices were to evolve differently from the subjective probabilities of a Bayesian agent, then the common interpretation of prediction market prices as “probabilities,” in the Bayesian sense, would be faulty. This makes the correct interpretation of prediction market prices as much an empirical question as a philosophical one.

We derive empirically-detectable features that prediction markets would exhibit if their contract prices were subjective probabilities. Then, using data from the Intrade Archive, we analyze the extent to which real-world prediction markets exhibit those features.

Literature Review

Prediction markets are a type of financial market where traders buy and sell contracts that function as bets on future events. In their simplest form, these contracts are converted into a single dollar if a certain prespecified event occurs and become worthless if the event does not occur (Oliven and Reitz, 2004). Intuitively, the price at which a contract trades should have a positive relationship with the likelihood of its associated event. Market participants who assign a higher probability to an event should be willing to bid more for contracts whose payout hinges on that event. Conversely, cheaper contracts should indicate that market participants view the event as less likely to occur. One could thus interpret prices in a prediction market as “probabilities” if they behaved like the beliefs of a Bayesian agent evolving over time. These exhibit certain identifiable features—in particular, a Bayesian agent’s beliefs should be a martingale with respect to the evidence uncovered at each time step (Aldous, 2013). It is possible to check whether this is true at various times before contracts’ maturity dates. However, the literature to date has focused on different considerations.

One such consideration is whether prediction markets produce probability estimates synthesized from the information held by market participants. In one of the earliest published papers on the Iowa Electronic Markets—the very first application of a prediction market mechanism—Forsythe, Nelson, Neumann, and Wright (1992) find that the market outperformed opinion polls in forecasting the outcome of the 1988 presidential election, and hypothesize that this is only due to the existence of a subset of traders who are free of judgment bias. The fact that many traders didn't possess this property did not seem to make the Iowa Electronic Markets less effective. Because of this, Forsythe et al. (1992) argue their findings support what Smith (1982) terms the "Hayek Hypothesis," which is the idea that markets can function well even if the participants have very limited knowledge. Many participants may be ignorant, but the market's "beliefs" can nevertheless be accurate.

This separation between the beliefs of a market's participants and the "beliefs" of the market itself—with the latter being represented by the price of a contract in the prediction market—is commonplace in the literature. Wolfers and Zitzewitz (2004) provide an introductory overview of prediction markets and explicitly describe the price of a contract as equivalent to the probability of its associated event, suggesting that such probabilities will be highly accurate even with very few market participants.

However, others disagree. Manski (2006), in one of the first formal analyses of price determination in a prediction market, constructs a model of prediction markets where the contract price is not the mean probability which participating traders would assign to the event's occurrence, but nevertheless yields a bound on the mean probability. He also finds that under this model, traders' knowledge of the current contract price does not change the equilibrium price. If this analysis is correct, it would be a mistake to conflate the contract price with the probability of a future event. Wolfers and Zitzewitz (2006) retort by producing a model where logarithmic utility yields a contract price equivalent to the mean belief among traders. While Forsythe et al. (1992) focuses on the problem of information synthesis, these debates are focused on the problem of belief aggregation. According to those on either side, if traders assign subjective probabilities to events, then contracts whose payouts are dependent on those events will have a price interpretable as a probability if and only if that price is a successful aggregation of the traders' subjective probabilities.

Another approach is to ignore the beliefs of market participants altogether. If the market itself assigns subjective probabilities to outcomes, then knowledge of past prices shouldn't allow one to produce a more accurate estimate of the chance of a future event than the present price alone. There have been some attempts to test this idea, but they are few in number. Berg, Forsythe, Nelson and Reitz (2008) summarize evidence from the Iowa Electronic Markets, showing that they yield predictions more accurate than those of large-scale polling organizations, which would suggest that knowledge of external information such as polls is already contained in the contract price. Additionally, Leigh, Wolfers, and Zitzewitz (2003) find that prediction market prices do not follow an easily predictable path, in that simple betting strategies based on past prices don't appear to yield opportunities for profit.

Berg and Rietz (2002), on the other hand, find less optimistic results. Their analysis of the data from the Iowa Electronic Markets demonstrates that traders are vulnerable to systematic overestimation and underestimation of the equilibrium contract price in response to sudden new private information. This allows for trading strategies that can exploit these biases when they occur. However, such trading strategies remain simple, and can only exploit biases that have already been detected. They do not come equipped with a way to determine when they should be employed. Additional research by Huber and Hauser (2005) finds that smaller prediction market contracts in Europe are systematically overpriced, which results in their failure to outperform polls as is the case in the United States. However, this finding does not extend to larger prediction market contracts and does not provide a concrete way that an observer could correct for the bias in general.

Papers which go any further in trying to analyze the consistency of prediction market contract prices with the behavior of Bayesian subjective probabilities are practically nonexistent. Tziralis and Tatsiopoulos (2012) conduct a literature review of the entire field from 1990 to 2006, and classify all 155 papers published during that period into distinct categories. A mere 26 are focused on the theoretical aspects of prediction markets and, of those, only two are dedicated to the interpretation of contract prices. A second literature review by Horn, Ivens, Ohneberg, and Brem (2014) performs the same classification on all 318 articles published between 2007 and 2013 and finds that only 57 are theoretical works. A search through the listed articles reveals that only one is dedicated to the interpretation of prices, where Keller, Mai, and Kros (2011) claim that prediction market prices are different from probabilities. They argue this can be explained by non-risk-neutral utility preferences and are more concerned with how individuals' subjective probabilities can be aggregated than with analyzing whether a market's beliefs evolve in a Bayesian manner.

These two literature reviews suggest that very little prior work has been done on modelling prediction markets as Bayesian agents. A rare exception is the research of Page and Clemens (2012), who analyze the calibration of prediction markets. They find that prediction markets are reasonably well calibrated when a contract's maturity date draws near, but fall prey to significant biases for contracts which will be resolved far in the future. This analysis is a special case of testing whether contract prices are martingales with respect to the evidence uncovered at each time step—a uniquely Bayesian characteristic. There is a distinct hole in the literature to be filled.

Theory

A martingale is a stochastic process where the conditional expectation of future values, given all previous values, is equal to the present value. A Bayesian agent believes that the subjective probabilities it assigns to a particular hypothesis will evolve through time like a martingale. Equivalently, an agent that does not believe its own beliefs will evolve through time like a martingale is not Bayesian.

Suppose an external observer took a record of how an agent updated its beliefs in response to several sequences of evidence, but was prohibited from viewing that evidence themselves. If the external observer were to detect empirical regularities in this record that are inconsistent with the agent believing its posterior will evolve like a martingale, he or she could determine that the agent is non-Bayesian. This same analysis applies to prediction markets. If their price data produce the same kinds of empirical regularities, it would be inaccurate to describe them as Bayesian agents.

It is important to distinguish between martingales with respect to themselves, and martingales with respect to another process. Let Ω denote a sample space. Then, a stochastic process $X : \mathbb{Z} \times \Omega \to \mathbb{R}$ is an integer-time martingale with respect to itself if, for all $n \in \mathbb{Z}$, the following properties hold:

- $\mathbb{E}(|X_n|) < \infty$
- $\mathbb{E}(X_{n+1} \mid X_n, X_{n-1}, ...) = X_n$.

Now let S be a set. A stochastic process $X : \mathbb{Z} \times \Omega \to \mathbb{R}$ is an integer-time martingale with respect to another stochastic process $Y : \mathbb{Z} \times \Omega \to S$ if, for all $n \in \mathbb{Z}$, the following properties hold:

- $\mathbb{E}(|X_n|) < \infty$
- $\mathbb{E}(X_{n+1} \mid Y_n, Y_{n-1}, ...) = X_n$.

Consider a Bayesian agent with prior I that observes evidence E_i at each discrete time step $i \in \mathbb{Z}$. For a hypothesis h, the Bayesian agent's posterior probability of h with respect to the evidence observed up until time t is

$$\mathbb{P}(h \mid I, E_t, E_{t-1}, ...).$$

Denoting the probability measure induced by the prior as $\mathbb{P}_I$ and the intersection all evidence up to time t as $\bigwedge_{i \leq t} E_i$, this becomes

$$\mathbb{P}_I(h \mid \bigwedge_{i \leq t} E_i).$$

Indexing these posterior probabilities to the integers produces a martingale with respect to the evidence so far observed by the Bayesian agent. The following properties thus hold:

- $\mathbb{E}_I(|\mathbb{P}_I(h \mid \bigwedge_{i\leq t} E_i)|) < \infty$
- $\mathbb{E}_I(\mathbb{P}_I(h \mid \bigwedge_{i\leq n+1} E_i) \mid \bigwedge_{i\leq n} E_i) = \mathbb{P}_I(h \mid \bigwedge_{i\leq n} E_i)$.

The first property follows directly from the fact that all probabilities fall between 0 and 1. The proof of the second martingale property is given below. For any $n \in \mathbb{Z}$, we have

$$\begin{aligned}
\mathbb{E}_I(\mathbb{P}_I(h \mid \bigwedge_{i\leq n+1} E_i) \mid \bigwedge_{i\leq n} E_i) &= \int_{E_{n+1}} \mathbb{P}_I(h \mid E_{n+1}, \bigwedge_{i\leq n} E_i)\mathbb{P}_I(E_{n+1} \mid \bigwedge_{i\leq n} E_i) \\
&= \int_{E_{n+1}} \mathbb{P}_I(h \wedge E_{n+1} \mid \bigwedge_{i\leq n} E_i) \\
&= \mathbb{P}_I(h \mid \bigwedge_{i\leq n} E_i).
\end{aligned}$$

This demonstrates that Bayesian agents believe their own posteriors will evolve through time as a martingale. Furthermore, by a property of integer-time martingales, the above equation holds for time steps greater than 1 into the future. For any $k \in \mathbb{Z}^+$,

$$\mathbb{E}_I(\mathbb{P}_I(h \mid \bigwedge_{i\leq n+k} E_i) \mid \bigwedge_{i\leq n} E_i) = \mathbb{P}_I(h \mid \bigwedge_{i\leq n} E_i).$$

For an external observer, ignorant of both any elements of the process $\{E_n\}$ or the prior I, this fact isn't very useful on its own. Fortunately, the external observer can use the martingale properties to derive detectable features of a Bayesian agent's updating process.

First, consider that the probability $\mathbb{P}(h \mid I, \bigwedge_{i\leq t} E_i)$ is a (potentially non-injective) function of the intersection of I and all evidence up to time t. Denote this function f, so that

$$f(I \cap (\bigcap_{i\leq t} E_i)) = \mathbb{P}(h \mid I, \bigwedge_{i\leq t} E_i).$$

By a property of images and preimages, we have

$$I \cap (\bigcap_{i\leq t} E_i) \subseteq f^{-1}(f(I \cap (\bigcap_{i\leq t} E_i)))$$

for any prior and any evidence observed. Therefore,

$$\begin{aligned}
&\mathbb{E}(\mathbb{P}(h \mid I, \bigwedge_{i\leq n+k} E_i) \mid I, \bigwedge_{i\leq n} E_i) \\
&= \mathbb{E}(\mathbb{P}(h \mid I, \bigwedge_{i\leq n+k} E_i) \mid I, \bigwedge_{i\leq n} E_i, \bigwedge_{i\leq n} \mathbb{P}(h \mid I, \bigwedge_{j\leq i} E_j)).
\end{aligned}$$

No new information is gained by conditioning on the probability of previous posteriors over and above that of the current evidence and prior. Substituting the above into the second martingale property, this becomes

$$\mathbb{E}(\mathbb{P}(h \mid I, \bigwedge_{i \leq n+k} E_i) \mid I, \bigwedge_{i \leq n} E_i, \bigwedge_{i \leq n} \mathbb{P}(h \mid I, \bigwedge_{j \leq i} E_j)) = \mathbb{P}(h \mid I, \bigwedge_{i \leq n} E_i).$$

Then, subtracting $\mathbb{P}(h \mid I, \bigwedge_{i \leq n} E_i)$ from both sides,

$$\mathbb{E}(\mathbb{P}(h \mid I, \bigwedge_{i \leq n+k} E_i) - \mathbb{P}(h \mid I, \bigwedge_{i \leq n} E_i) \mid I, \bigwedge_{i \leq n} E_i, \bigwedge_{i \leq n} \mathbb{P}(h \mid I, \bigwedge_{j \leq i} E_j)) = 0.$$

Suppose now that the external observer selects some finite set of previously observed probabilities from the agent, all from at or before time period n. Denote this set D. Then, applying the conditional expectation $\mathbb{E}(\cdot \mid D)$ to both sides of the above equation,

$$\mathbb{E}(\mathbb{E}(\mathbb{P}(h \mid I, \bigwedge_{i \leq n+k} E_i) - \mathbb{P}(h \mid I, \bigwedge_{i \leq n} E_i) \mid I, \bigwedge_{i \leq n} E_i, \bigwedge_{i \leq n} \mathbb{P}(h \mid I, \bigwedge_{j \leq i} E_j)) \mid D) = 0$$

and by the law of iterated expectations,

$$\mathbb{E}(\mathbb{P}(h \mid I, \bigwedge_{i \leq n+k} E_i) - \mathbb{P}(h \mid I, \bigwedge_{i \leq n} E_i) \mid D) = 0. \tag{1}$$

The external observer can make use of this above expression to test whether a prediction market is a Bayesian agent, since the expectation is taken with respect to a set of observable prices D. He or she need not know the evidence upon which the prediction market has conditioned, nor its prior.

If, for all times that a set of prices D are observed, the difference between any two prediction market prices some fixed distance from one another which both occur chronologically after D has a mean of 0, then the prediction market data would be consistent with the behavior of Bayesian updating. A special case of this idea is calibration testing, where the distance is taken between a contract's resolving price (either 1 or 0) and its price at some fixed distance in the past.

Let C denote the set of all contracts, let $p_{c,i}$ denote the price of contract c at time i away from the maturity date, and let q_d denote the time between the contract price evaluated at $d \in D$ and at time n. Lastly, let T_D denote the difference between the time of the first and last price in D. To compute an estimate of the left-hand side of equation (1) we use a kernel type estimator, with beta kernels (denoted as Beta below) set to a fixed concentration parameter and centered at each point in D:

$$\frac{\sum_{c\in C}\sum_{i=1}^{|c|-k-T_D}(p_{c,i}-p_{c,i+k})\prod_{d\in D}\mathrm{Beta}(p_{c,i+k+q_d},d)}{\sum_{c\in C}\sum_{i=1}^{|c|-k-T_D}\prod_{d\in D}\mathrm{Beta}(p_{c,i+k+q_d},d)}$$

If prediction markets are not Bayesian agents, there should exist some choice of D, q, and k such that the above estimator is distant from zero.

Evidence

Intrade, one of the world's largest prediction markets, was shut down in 2013. Fortunately, Intrade's market data were preserved for posterity. The Intrade Archive is a collection of time series that correspond to individual prediction contracts, which contain crucial market information recorded over a contract's lifetime (e.g. closing prices, trading volume per day). Panos Ipeirotis, the creator of the Intrade Archive and a prediction market researcher, aggregated the data from all non-financial contracts into a single repository and made them available to the public. It is these data which will be used to test whether prediction markets behave as Bayesian agents.

The 17287 different time series in the repository originate between 2003 and 2013. Not all are usable. The closing prices for each day are quoted in cents, and so their values should range between 0 and 100. Some of the time series, however, display pathological behavior, and contain contracts traded at least once for a price greater than a dollar. Only 9 time series display this unusual behavior. For the purpose of this analysis, these time series were removed.

A larger problem comes in the form of missing price data. 3378 time series possess empty entries for their contract's closing price at certain dates. This is a common difficulty encountered in historical financial data. It is possible to interpolate between recorded closing prices within the affected time series, but for the purpose of this analysis they were removed entirely. Additionally, there is significant overlap between different kinds of pathological time series–all those with closing prices greater than a dollar also possessed incomplete closing price data.

Lastly, some contracts were resolved or cancelled the same day they were issued, and thus their associated time series have a length of zero. 6 contracts displayed this property. The window in kernel regression has a finite, non-zero length, and so will ignore these time series altogether. For this reason, zero-length time series were retained, but will have no impact on the results of the analysis.

Removed time series tended to be heavily composed of missing or irregular data; the mean percentage of missing or irregular data points in a time series that contained missing or irregular data was 49%. This suggests that missing data are clustered among a handful of files, rather than spread out among the entire repository.

After partitioning the repository into time series with incomplete or pathological data and those without, summary statistics were collected for each. The results are as follows:

DATA SET	All time series	Irregular/removed	Retained
Number of time series	17287	3378	13909
Average length	117 days	153 days	108 days
Standard Deviation of Length	212 days	254 days	199 days
Minimum Length	0 days	1 day	0 days
Maximum Length	1937 days	1937 days	1886 days

The beta kernel accepts values between 0 and 1 as inputs. Since closing prices are quoted in cents (and thus lie between 0 and 100), they were converted into fractions of a dollar to be compatible with the beta kernel. No other pre-processing was performed on the data.

Data Analysis

Prediction markets do not violate the properties of a Bayesian agent if equation (1) holds. Testing whether this is the case requires computing an estimate of the expected value and the variance of

$$[\mathbb{P}(h \mid I, \bigwedge_{i \leq n+k} E_i) - \mathbb{P}(h \mid I, \bigwedge_{i \leq n} E_i) \mid D]. \tag{2}$$

An estimate of the expected value of (2) can be computed from the data as

$$\mu = \frac{\sum_{c\in C}\sum_{i=1}^{|c|-k-T_D}(p_{c,i}-p_{c,i+k})\prod_{d\in D}\mathrm{Beta}(p_{c,i+k+q_d},d)}{\sum_{c\in C}\sum_{i=1}^{|c|-k-T_D}\prod_{d\in D}\mathrm{Beta}(p_{c,i+k+q_d},d)},$$

and an estimate of the variance of (2) can be similarly computed as

$$\sigma^2 = \frac{\sum_{c\in C}\sum_{i=1}^{|c|-k-T_D}(p_{c,i}-p_{c,i+k})^2\prod_{d\in D}\mathrm{Beta}(p_{c,i+k+q_d},d)}{\sum_{c\in C}\sum_{i=1}^{|c|-k-T_D}\prod_{d\in D}\mathrm{Beta}(p_{c,i+k+q_d},d)} - \mu^2.$$

This requires the specification of values for k, q, and D, where k is the distance between the two prices whose average difference is being tested, q is the distance between the earlier of these two prices and the latest price in D, and D is the set of prices conditioned upon.

Additionally, the above beta kernels utilize the mode-concentration parameterization. For this analysis, a high concentration of 30 was selected, so as to better avoid conditioning upon prices too distant from those in D. Each mode is set to the value of d (the price conditioned upon).

The simplest test of whether (1) holds would be to compute the expected deviation between a current price and the next day's closing price, which corresponds to $k = 1$ and $q = 0$. D will consist of all possible prices between 0 and 100 cents at 5-cent intervals (i.e. 0, 0.05, 0.10, etc.). Due to the nature of kernel regression, it is not possible to determine the size of the data set used to compute these expected deviations, and so t-tests cannot be legitimately applied to the result. As a substitute, each estimate is accompanied by a signal to noise ratio:

$$\frac{\mu}{\sigma}$$

μ and σ refer to our point estimate of the expected value of (2) and its standard deviation. A high signal to noise ratio indicates clear and predictable deviation of (2) from zero, while a low signal to noise ratio indicates that either no such deviations can be found or that deviations are difficult to predict.

The results of this test are displayed in the following two columns. The first contains the point estimates μ for different starting prices, and the second contains the corresponding signal to noise ratios.

D	μ	$\frac{\mu}{\sigma}$
0.0	0.01	0.1311
0.05	0.0041	0.0804
0.1	0.0016	0.0387
0.15	-0.0006	-0.0163
0.2	-0.002	-0.055
0.25	-0.0031	-0.0717
0.3	-0.004	-0.0789
0.35	-0.0048	-0.0829
0.4	-0.0057	-0.0872
0.45	-0.0066	-0.0922
0.5	-0.0074	-0.0966
0.55	-0.0079	-0.0989
0.6	-0.0081	-0.0993
0.65	-0.0083	-0.0992
0.7	-0.0085	-0.0998
0.75	-0.0089	-0.1015
0.8	-0.0094	-0.1038
0.85	-0.01	-0.1066
0.9	-0.0109	-0.1101
0.95	-0.0119	-0.1147
1.0	-0.0139	-0.1188

These results suggest that there is a small but nearly insignificant negative bias in Intrade prediction markets over the extremely short term. None of the signal to noise ratios are large enough to demonstrate inconsistency with the behavior of a Bayesian agent.

A second and more sophisticated test would involve increasing the size of k or q. Suppose two prediction market prices observed a week apart. What bias might result in price movements a month out ($k = 28$)?

The results of this test are displayed in two separate tables. The first corresponds to the values of μ, and the second corresponds to the values of the signal to noise ratio. Prices at the start of the week are plotted on the x-axis, and prices at the end of the week are plotted on the y-axis.

D	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.0	0.0285	0.0301	0.0697	0.1809	0.2755	0.3475	0.4406	0.5765	0.6827	0.7089	0.75
0.1	0.0241	-0.0006	-0.0041	-0.0033	0.0134	0.077	0.1886	0.3624	0.536	0.6717	0.75
0.2	0.0009	-0.0065	-0.0097	-0.0118	-0.013	-0.01	0.003	0.0278	0.0919	0.2744	0.75
0.3	-0.0301	-0.0129	-0.0131	-0.0141	-0.0145	-0.0142	-0.0147	-0.0199	-0.0151	0.0227	0.75
0.4	-0.055	-0.0309	-0.0171	-0.0154	-0.0146	-0.0142	-0.015	-0.0199	-0.0274	-0.0034	0.7498
0.5	-0.0753	-0.0709	-0.0275	-0.0184	-0.0153	-0.0143	-0.015	-0.0179	-0.022	-0.0122	0.0104
0.6	-0.0762	-0.0892	-0.0612	-0.0268	-0.0175	-0.0156	-0.016	-0.0175	-0.019	-0.014	-0.0217
0.7	-0.0604	-0.0784	-0.1094	-0.05	-0.0267	-0.0199	-0.0183	-0.0184	-0.018	-0.0153	-0.0217
0.8	-0.0564	-0.0687	-0.138	-0.1051	-0.0753	-0.0403	-0.0246	-0.0197	-0.017	-0.0156	-0.0217
0.9	-0.0792	-0.0841	-0.1497	-0.1585	-0.1961	-0.1786	-0.0573	-0.0227	-0.0177	-0.0164	-0.0217
1.0	-0.0909	-0.0909	-0.0909	-0.0909	-0.0909	-0.0256	-0.0217	-0.0217	-0.0217	-0.0217	-0.0217

D	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.0	0.2217	0.2401	0.5239	1.035	1.1812	1.2725	1.4862	1.9144	2.1247	1.8967	1.7321
0.1	0.1955	-0.0106	-0.0623	-0.0391	0.1002	0.3485	0.62	0.9816	1.3404	1.69	1.7321
0.2	0.0075	-0.0927	-0.1184	-0.1251	-0.1194	-0.076	0.0176	0.1243	0.3009	0.6241	1.7321
0.3	-0.1968	-0.1351	-0.133	-0.1311	-0.1237	-0.1103	-0.1005	-0.1104	-0.0642	0.0762	1.7321
0.4	-0.2808	-0.2092	-0.1423	-0.1277	-0.1176	-0.1105	-0.1102	-0.1282	-0.1323	-0.0121	1.731
0.5	-0.3397	-0.3286	-0.1751	-0.1331	-0.1157	-0.1094	-0.1127	-0.1256	-0.1301	-0.0537	0.0456
0.6	-0.3379	-0.3801	-0.2875	-0.1612	-0.1228	-0.1156	-0.1199	-0.1266	-0.1275	-0.0827	-0.1489
0.7	-0.2752	-0.3369	-0.4377	-0.2377	-0.161	-0.1369	-0.1317	-0.1326	-0.13	-0.1123	-0.1489
0.8	-0.2519	-0.2943	-0.5231	-0.4046	-0.3157	-0.2156	-0.1609	-0.1426	-0.1313	-0.1263	-0.1489
0.9	-0.3123	-0.3297	-0.5494	-0.506	-0.5454	-0.5104	-0.2703	-0.1644	-0.142	-0.1365	-0.1489
1.0	-0.3162	-0.3162	-0.3162	-0.3162	-0.3162	-0.162	-0.1489	-0.1489	-0.1489	-0.1489	-0.1489

The results of this test show stronger violations of equation (1). As prices at the start and end of the week drift toward either 0 or 1, clear biases start to emerge. The bias is minor if both prices are near 0 or 1, and grows stronger if one price is near 0 and the other near 1. The former situation should occur if a prediction market is "almost certain" about its beliefs, and fails to encounter contrary evidence over the course of a week. The latter situation should occur if, within that week, a prediction market receives highly unexpected evidence that causes it to sharply reevaluate its own beliefs.

Conclusion

Prediction markets appear to behave the least like Bayesian agents when updating in response to unexpected evidence, or when they grow strongly confident in a specific conclusion. Prediction markets' behavior is more consistent with that of a Bayesian agent when neither of these conditions hold.

The methods used here are not intended as a substitute for calibration analysis. As a consequence, the above results are not a strong indicator that prediction markets become unreliable forecasters when the conditions for non-Bayesian behavior occur. However, they should caution against interpreting prediction market prices as the probabilistic beliefs of a Bayesian agent in such circumstances.

References

Smith, V. L. (1982). Markets as economizers of information: experimental examination of the “Hayek hypothesis”. *Economic Inquiry*, *20*(2), 165-179.

Forsythe, R., Nelson, F., Neumann, G. R., Wright, J. (1992). Anatomy of an experimental political stock market. *The American Economic Review*, 1142-1161.

Leigh, A., Wolfers, J., Zitzewitz, E. (2003). *What do financial markets think of war in Iraq?* (No. w9587). National Bureau of Economic Research.

Oliven, K., Rietz, T. A. (2004). Suckers are born but markets are made: Individual rationality, arbitrage, and market efficiency on an electronic futures market. *Management Science*, *50*(3), 336-351.

Wolfers, J., Zitzewitz, E. (2004). Prediction markets. *Journal of economic perspectives*, *18*(2), 107-126.

Huber, J., Hauser, F. (2005). Systematic mispricing in experimental markets evidence from political stock markets. In *Proceedings of the International Conference on Finance, Kopenhagen, Denmark*.

Wolfers, J., Zitzewitz, E. (2006). *Interpreting prediction market prices as probabilities* (No. w12200). National Bureau of Economic Research.

Manski, C. F. (2006). Interpreting the predictions of prediction markets. *economics letters*, *91*(3), 425-429.

Berg, J., Forsythe, R., Nelson, F., Rietz, T. (2008). Results from a dozen years of election futures markets research. *Handbook of experimental economics results*, *1*, 742-751.

Keller, C. M., Mai, E., Kros, J. F. (2011). Interpreting Political Prediction Market Prices as Probabilities: A Study of a 2008 U.S. Presidential Election Market. *Journal of Prediction Markets*, *5*(3).

Page, L., Clemen, R. T. (2012). Do Prediction Markets Produce Well-Calibrated Probability Forecasts?. *The Economic Journal*, *123*(568), 491-513.

Tziralis, G., Tatsiopoulos, I. (2012). Prediction markets: An extended literature review. *The journal of prediction markets*, *1*(1), 75-91.

Aldous, D. J. (2013). Using prediction market data to illustrate undergraduate probability. *The American Mathematical Monthly*, *120*(7), 583-593.

Horn, C. F., Ivens, B. S., Ohneberg, M., Brem, A. (2014). Prediction markets-a literature review 2014. *The journal of prediction markets*, *8*(2), 89-126.