

Prediction Markets: Tales from the Election

Special thanks to Jeff Coleman, Karl Floersch and Robin Hanson for critical feedback and review.

Trigger warning: I express some political opinions.

Prediction markets are a subject that has interested me for many years. The idea of allowing anyone in the public to make bets about future events, and using the odds at which these bets are made as a [credibly neutral](#) source of predicted probabilities of these events, is a fascinating application of mechanism design. Closely related ideas, like [futarchy](#), have always interested me as innovative tools that could improve governance and decision-making. And as [Augur](#) and [Omen](#), and more recently [PolyMarket](#), have shown, prediction markets are a fascinating application of blockchains (in all three cases, Ethereum) as well.

And the 2020 US presidential election, it seems like prediction markets are finally entering the limelight, with blockchain-based markets in particular growing from near-zero in 2016 to [millions of dollars of volume in 2020](#). As someone who is closely interested in seeing Ethereum applications cross the chasm into widespread adoption, this of course aroused my interest. At first, I was inclined to simply watch, and not participate myself: I am not an expert on US electoral politics, so why should I expect my opinion to be more correct than that of everyone else who was already trading? But in my Twitter-sphere, I saw more and more arguments from Very Smart People whom I respected arguing that the markets were

in fact being irrational and I should

participate and bet against them if I can. Eventually, I was convinced.

I decided to make an experiment on the blockchain that I helped to create: I bought \$2,000 worth of NTRUMP (tokens that pay \$1 if Trump loses) on Augur. Little did I know then that my position would eventually increase to \$308,249, earning me a profit of over \$56,803, and that I would make all of these remaining bets, against willing counterparties, after Trump had already lost the election

. What would transpire over the next two months would prove to be a fascinating case study in social psychology, expertise, arbitrage, and the limits of market efficiency, with important ramifications to anyone who is deeply interested in the possibilities of economic institution design.

Before the Election

My first bet on this election was actually not on a blockchain at all. When [Kanye announced his presidential bid](#) in July, a political theorist whom I ordinarily quite respect for his high-quality and original thinking immediately claimed on Twitter that he was confident that this would split the anti-Trump vote and lead to a Trump victory. I remember thinking at the time that this particular opinion of his was over-confident, perhaps even a result of over-internalizing the heuristic that if a viewpoint seems clever and contrarian then it is likely to be correct. So [of course](#) I offered to make a \$200 bet, myself betting the boring conventional pro-Biden view, and he honorably accepted.

The election came up again on my radar in September, and this time it was the prediction markets that caught my attention. The markets gave Trump a nearly 50% chance of winning, but I saw many Very Smart People in my Twitter-sphere whom I respected pointing out that this number seemed far too high. This of course led to the familiar "efficient markets debate": if you can buy a token that gives you \$1 if Trump loses for \$0.52, and Trump's actual chance of losing is much higher, why wouldn't people just come in and buy the token until the price rises more? And if nobody has done this, who are you to think that you're smarter than everyone else?

Neoliberal's [Twitter thread just before Election Day](#) does an excellent job summarizing his case against prediction markets being accurate at that time. In short, the (non-blockchain) prediction markets that most people used at least prior to 2020 have all sorts of restrictions that make it difficult for people to participate with more than a small amount of cash. As a result, if a very smart individual or a professional organization saw a probability that they believed was wrong, they would only have a very limited ability to push the price in the direction that they believe to be correct.

The most important restrictions that the paper points out are:

- Low limits (well under \$1,000) on how much each person can bet
- High fees (eg. a 5% withdrawal fee on PredictIt)

And this is where I [pushed back against neoliberal in September](#): although the stodgy old-world centralized prediction markets may have low limits and high fees, the crypto markets do not! On Augur or Omen, there's no limit to how much someone can buy or sell if they think the price of some outcome token is too low or too high. And the blockchain-based prediction markets were following the same prices as PredictIt. If the markets really were over-estimating Trump because high fees and low trading limits were preventing the more cool-headed traders from outbidding the overly optimistic ones, then why would blockchain-based markets, which don't have those issues, show the same prices?

The main response my Twitter friends gave to this was that blockchain-based markets are highly niche, and very few people, particularly very few people who know much about politics, have easy access to cryptocurrency. That seemed plausible, but I was not too

confident in that argument. And so I bet \$2,000 against Trump and went no further.

The Election

Then the election happened. After an initial scare where Trump at first won more seats than we expected, Biden turned out to be the eventual winner. Whether or not the election itself validated or refuted the efficiency of prediction markets is a topic that, as far as I can tell, is quite open to interpretation. On the one hand, by a standard [Bayes rule](#) application, I should decrease my confidence of prediction markets, at least relative to Nate Silver. Prediction markets gave a 60% chance of Biden winning, Nate Silver gave a [90% chance of Biden winning](#). Since Biden in fact won, this is one piece of evidence that I live in a world where Nate gives the more correct answers.

But on the other hand, you can make a case that the prediction markets better estimated the margin

of victory. The median of Nate's probability distribution was somewhere around 370 of 538 electoral college votes going to Biden:

The Trump markets didn't give a probability distribution, but if you had to

guess a probability distribution from the statistic "40% chance Trump will win", you would probably give one with a median somewhere around 300 EC votes for Biden. The actual result: 306. So the net score for prediction markets vs Nate seems to me, on reflection, ambiguous.

After the election

But what I could not have imagined at the time was that the election itself was just the beginning. A few days after the election, Biden was declared the winner by various major organizations and even a few foreign governments. Trump mounted various legal challenges to the election results, as was expected, but each of these challenges [quickly failed](#). But for over a month, the price of the NTRUMP tokens stayed at 85 cents

!

At the beginning, it seemed reasonable to guess that Trump had a 15% chance of overturning the results; after all, he had appointed [three judges](#) to the Supreme Court, at a time of heightened partisanship where many have come to favor team over principle. Over the next three weeks, however, it became more and more clear that the challenges were failing, and Trump's hopes continued to look grimmer with each passing day, but the NTRUMP price did not budge; in fact, it even briefly decreased

to around \$0.82. On December 11, more than five weeks after the election, the [Supreme Court decisively and unanimously rejected](#) Trump's attempts to overturn the vote, and the NTRUMP price finally rose.... to \$0.88.

It was in November that I was finally convinced that the market skeptics were right, and I plunged in and bet against Trump myself. The decision was not so much about the money; after all, barely two months later I would [earn and donate to GiveDirectly](#) a far larger amount simply from holding dogecoin. Rather, it was to take part in the experiment not just as an

observer, but as an active participant, and improve my personal understanding of why everyone else hadn't already plunged in to buy NTRUMP tokens before me.

Dipping in

I bought my NTRUMP on [Catnip](#), a front-end user interface that combines together the Augur prediction market with [Balancer](#), a [Uniswap-style constant-function market maker](#). Catnip was by far the easiest interface for making these trades, and in my opinion contributed significantly to Augur's usability.

There are two ways to bet against Trump with Catnip:

1. Use [DAI](#) to buy NTRUMP on Catnip directly
2. Use [Foundry](#) to access an Augur feature that allows you to convert 1 DAI into 1 NTRUMP + 1 YTUMP + 1 ITRUMP (the "I" stands for "invalid", more on this later), and sell the YTRUMP on Catnip

At first, I only knew about the first option. But then I discovered that Balancer has far more liquidity for YTRUMP, and so I switched to the second option.

There was also another problem: I did not have any DAI. I had ETH, and I could have sold my ETH to get DAI, but I did not want to sacrifice my ETH exposure; it would have been a shame if I earned \$50,000 betting against Trump but simultaneously lost \$500,000 missing out on ETH price changes. So I decided to keep my ETH price exposure the same by opening up a collateralized debt position (CDP, now also called a ["vault"](#)) on MakerDAO.

A CDP is how all DAI is generated: users deposit their ETH into a smart contract, and are allowed to withdraw an amount of newly-generated DAI up to 2/3 of the value of ETH that they put in. They can get their ETH back by sending back the same amount of DAI that they withdrew plus an extra interest fee (currently 3.5%). If the value of the ETH collateral that you deposited drops to less than 150% the value of the DAI you withdrew, anyone can come in and ["liquidate"](#) the vault, forcibly selling the ETH to buy back the DAI and charging you a high penalty. Hence, it's a good idea to have a high collateralization ratio in case of sudden price movements; I had over \$3 worth of ETH in my CDP for every \$1 that I withdrew.

Recapping the above, here's the pipeline in diagram form:

I did this many times; the slippage on Catnip meant that I could normally make trades only up to about \$5,000 to \$10,000 at a time without prices becoming too unfavorable (when I had skipped Foundry and bought NTRUMP with DAI directly, the limit was closer to \$1,000). And after two months, I had accumulated over 367,000 NTRUMP.

Why not everyone else?

Before I went in, I had four main hypotheses about why so few others were buying up dollars for 85 cents:

1. Fear that either the Augur smart contracts would break or Trump supporters would manipulate the oracle (a [decentralized mechanism](#) where holders of Augur's REP token vote by staking their tokens on one outcome or the other) to make it return a false result
2. Capital costs: to buy these tokens, you have to lock up funds for over two months, and this removes your ability to spend those funds or make other profitable trades for that duration
3. It's too technically complicated for almost everyone to trade
4. There just really are far fewer people than I thought who are actually motivated enough to take a weird opportunity even when it presents them straight in the face

All four have reasonable arguments going for them. Smart contracts breaking [is a real risk](#), and the Augur oracle had not before been tested in such a contentious environment. Capital costs are real, and while betting against something is easier in a prediction market than in a stock market because you know that prices will never go above \$1, locking up capital nevertheless competes with other lucrative opportunities in the crypto markets. Making transactions things in dapps is technically complicated, and it's rational to have some degree of fear-of-the-unknown.

But my experience actually going into the financial trenches, and watching the prices on this market evolve, taught me a lot about each of these hypotheses.

Fear of smart contract exploits

At first, I thought that "fear of smart contract exploits" must have been a significant part of the explanation. But over time, I have become more convinced that it is probably not

a dominant factor. One way to see why I think this is the case is to compare the prices for YTRUMP and ITRUMP. ITRUMP stands for "Invalid Trump"; "Invalid" is an event outcome that is intended to be triggered in [some exceptional cases](#): when the description of the event is ambiguous, when the outcome of the event is not yet known when the market is resolved, when the market is unethical (eg. assassination markets), and a few other similar situations. In this market, the price of ITRUMP consistently stayed under \$0.02. If someone wanted to earn a profit by attacking the market, it would be far more lucrative for them to not buy YTRUMP at \$0.15, but instead buy ITRUMP at \$0.02. If they buy a large amount of ITRUMP, they could earn a 50x return if they can force the "invalid" outcome to actually trigger. So if you fear an attack, buying ITRUMP is by far the most rational response. And yet, very few people did.

A further argument against fear of smart contract exploits, of course, is the fact that in every crypto application except prediction markets (eg. Compound, the various yield farming schemes) people are surprisingly blasé about smart contract risks. If people are willing to put their money into all sorts of risky and untested schemes even for a promise of mere 5-8% annual gains, why would they suddenly become over-cautious here?

Capital costs

Capital costs - the inconvenience and opportunity cost of locking up large amounts of money - are a challenge that I have come to appreciate much more than I did before. Just looking at the Augur side of things, I needed to lock up 308,249 DAI for an average of about two months to make a \$56,803 profit. This works out to about a 175% annualized interest rate; so far, quite a good deal, even compared to the various yield farming crazes of the summer of 2020. But this becomes worse when you take into account what I needed to do on MakerDAO. Because I wanted to keep my exposure to ETH the same, I needed to get my DAI through a CDP, and safely using a CDP required a collateral ratio of over 3x. Hence, the total amount of capital I actually

needed to lock up was somewhere around a million dollars.

Now, the interest rates are looking less favorable. And if you add to that the possibility, however remote, that a smart contract hack, or a truly unprecedented political event, actually will

happen, it looks less favorable still.

But even still, assuming a 3x lockup and a 3% chance of Augur breaking (I had bought ITRUMP to cover the possibility that it breaks in the "invalid" direction, so I needed only worry about the risk of breaks in the "yes" direction or the the funds being stolen outright), that works out to a risk-neutral rate of about 35%, and even lower once you take real human beings' views on risk into account. The deal is still very attractive, but on the other hand, it now looks very understandable that such numbers are unimpressive to people who live and breathe cryptocurrency with its frequent 100x ups and downs.

Trump supporters

, on the other hand, faced none of these challenges: they cancelled out my \$308,249 bet by throwing in a mere \$60,000 (my winnings are less than this because of fees). When probabilities are close to 0 or 1, as is the case here, the game is very

lopsided in favor of those who are trying to push the probability away from the extreme value. And this explains not just Trump; it's also the reason why all sorts of popular-among-a-niche candidates with no real chance of victory frequently get winning probabilities as high as 5%.

Technical complexity

I had at first tried buying NTRUMP on Augur, but technical glitches in the user interface prevented me from being able to

make orders on Augur directly (other people I talked to did not have this issue... I am still not sure what happened there). Catnip's UI is much simpler and worked excellently. However, automated market makers like Balancer (and Uniswap) work best for smaller trades; for larger trades, the slippage is too high. This is a good microcosm of the broader "AMM vs order book" debate: AMMs are more convenient but order books really do work better for large trades. Uniswap v3 is introducing an AMM design that has better capital efficiency; we shall see if that improves things.

There were other technical complexities too, though fortunately they all seem to be easily solvable. There is no reason why an interface like Catnip could not integrate the "DAI -> Foundry -> sell YTRUMP" path into a contract so that you could buy NTRUMP that way in a single transaction. In fact, the interface could even check the price and liquidity properties of the "DAI -> NTRUMP" path and the "DAI -> Foundry -> sell YTRUMP" path and give you the better trade automatically. Even withdrawing DAI from a MakerDAO CDP can be included in that path. My conclusion here is optimistic: technical complexity issues were a real barrier to participation this round, but things will be much easier in future rounds as technology improves.

Intellectual underconfidence

And now we have the final possibility: that many people (and smart people in particular) have a pathology that they suffer from excessive humility, and too easily conclude that if no one else has taken some action, then there must therefore be a good reason why that action is not worth taking.

Eliezer Yudkowsky spends the second half of his excellent book [Inadequate Equilibria](#) making this case, arguing that too many people overuse "modest epistemology", and we should be much more willing to act on the results of our reasoning, even when the result suggests that the great majority of the population is irrational or lazy or wrong about something. When I read those sections for the first time, I was unconvinced; it seemed like Eliezer was simply being overly arrogant. But having gone through this experience, I have come to see some wisdom in his position.

This was not my first time seeing the virtues of trusting one's own reasoning first hand. When I had originally started working on Ethereum, I was at first beset by fear that there must be some very good reason the project was doomed to fail. A fully programmable smart-contract-capable blockchain, I reasoned, was clearly such a great improvement over what came before, that surely many other people must have thought of it before I did. And so I fully expected that, as soon as I publish the idea, many very smart cryptographers would tell me the very good reasons why something like Ethereum was fundamentally impossible. And yet, no one ever did.

Of course, not everyone suffers from excessive modesty. Many of the people making predictions in favor

of Trump winning the election were arguably fooled by their own excessive contrarianism. Ethereum benefited from my youthful suppression of my own modesty and fears, but there are plenty of other projects that could have benefited from more intellectual humility and avoided failures.

But nevertheless it seems to me more true than ever that, as goes the [famous Yeats quote](#), "the best lack all conviction, while the worst are full of passionate intensity." Whatever the faults of overconfidence or contrarianism sometimes may be, it seems clear to me that spreading a society-wide message that the solution is to simply trust the existing outputs of society, whether those come in the form of academic institutions, media, governments or markets, is not

the solution. All of these institutions can only work precisely because of the presence of individuals who think that they do not work, or who at least think that they can be wrong at least some of the time.

Lessons for futarchy

Seeing the importance of capital costs and their interplay with risks first hand is also important evidence for judging systems like [futarchy

](<https://blog.ethereum.org/2014/08/21/introduction-futarchy/>). Futarchy, and "decision markets" more generally are an important and potentially very socially useful application of prediction markets. There is not much social value in having slightly more accurate predictions of who will be the next president. But there is

a lot of social value in having conditional predictions

: if we do A, what's the chance it will lead to some good thing X, and if we do B instead what are the chances then

? Conditional predictions are important because they do not just satisfy our curiosity; they can also help us make decisions.

Though electoral prediction markets are much less useful than conditional predictions, they can help shed light on an important question: how robust are they to manipulation or even just biased and wrong opinions? We can answer this question by looking at how difficult arbitrage is: suppose that a conditional prediction market currently gives probabilities that (in your opinion) are wrong

(could be because of ill-informed traders or an explicit manipulation attempt; we don't really care). How much of an impact can you have, and how much profit can you make, by setting things right?

Let's start with a concrete example. Suppose that we are trying to use a prediction market to choose between decision A and decision B, where each decision has some probability of achieving some desirable outcome. Suppose that your opinion is that decision A has a 50% chance of achieving the goal, and decision B has a 45% chance. The market, however, (in your opinion wrongly) thinks decision B has a 55% chance and decision A has a 40% chance.

Suppose that you are a small participant, so your individual bets won't affect the outcome; only many bettors acting together could. How much of your money should you bet?

The standard theory here relies on the [Kelly criterion](#). Essentially, you should act to maximize the expected logarithm of your assets. In this case, we can solve the resulting equation. Suppose you invest portion (r)

of your money into buying A-token for \$0.4. Your expected new log-wealth, from your point of view, would be:

$$(0.5 * \log((1-r) + \frac{r}{0.4}) + 0.5 * \log(1-r))$$

The first term is the 50% chance (from your point of view) that the bet pays off, and the portion (r)

that you invest grows by 2.5x (as you bought dollars at 40 cents). The second term is the 50% chance that the bet does not pay off, and you lose the portion you bet. We can use calculus to find the (r)

that maximizes this; for the lazy, [here's WolframAlpha](#). The answer is ($r = \frac{1}{6}$)

. If other people buy and the price for A on the market gets up to 47% (and B gets down to 48%), we can redo the calculation for the last trader who would flip the market over to make it correctly favor A:

$$(0.5 * \log((1-r) + \frac{r}{0.47}) + 0.5 * \log(1-r))$$

Here, the expected-log-wealth-maximizing (r)

is a mere 0.0566. The conclusion is clear: when decisions are close and when there is a lot of noise, it turns out that it only makes sense to invest a small portion of your money in a market. And this is assuming rationality; most people invest less

into uncertain gambles than the Kelly criterion says they should. Capital costs stack on top even further. But if an attacker really

wants to force outcome B through because they want it to happen for personal reasons, they can simply put all

of their capital toward buying that token. All in all, the game can easily be lopsided more than 20:1 in favor of the attacker.

Of course, in reality attackers are rarely willing to stake all their funds on one decision. And futarchy is not the only mechanism that is vulnerable to attacks: stock markets are similarly vulnerable, and non-market decision mechanisms can also be manipulated by determined wealthy attackers in all sorts of ways. But nevertheless, we should be wary of assuming that futarchy will propel us to new heights of decision-making accuracy.

Interestingly enough, the math seems to suggest that futarchy would work best when the expected manipulators would want to push the outcome toward

an extreme value. An example of this might be liability insurance, as someone wishing to improperly obtain insurance would effectively be trying to force the market-estimated probability that an unfavorable event will happen down to zero. And as it turns out, liability insurance is futarchy inventor Robin Hanson's [new favorite policy prescription](#).

Can prediction markets become better?

The final question to ask is: are prediction markets doomed to repeat errors as grave as giving Trump a 15% chance of overturning the election in early December, and a 12% chance of overturning it even after the Supreme Court including three judges whom he appointed telling him to screw off? Or could the markets improve over time? My answer is, surprisingly, emphatically on the optimistic side, and I see a few reasons for optimism.

Markets as natural selection

First, these events have given me a new perspective on how market efficiency and rationality might actually come about. Too often, proponents of market efficiency theories claim that market efficiency results because most participants are rational (or at least the rationals outweigh any coherent group of deluded people), and this is true as an axiom. But instead, we could take an evolutionary

perspective on what is going on.

Crypto is a young ecosystem. It is an ecosystem that is still quite disconnected from the mainstream, Elon's recent tweets notwithstanding, and that does not yet have much expertise in the minutiae of electoral politics. Those who are

experts in electoral politics have a hard time getting into crypto, and crypto has a large presence of not-always-correct forms of contrarianism especially when it comes to politics. But what happened this year is that within the crypto space, prediction market users who correctly expected Biden to win got an 18% increase to their capital, and prediction market users who incorrectly expected Trump to win got a 100% decrease to their capital (or at least the portion they put into the bet).

Thus, there is a selection pressure

in favor of the type of people who make bets that turn out to be correct. After ten rounds of this, good predictors will have more capital to bet with, and bad predictors will have less capital to bet with. This does not

rely on anyone "getting wiser" or "learning their lesson" or any other assumption about humans' capacity to reason and learn. It is simply a result of selection dynamics that over time, participants that are good at making correct guesses will come to dominate the ecosystem.

Note that prediction markets fare better than stock markets in this regard: the "nouveau riche" of stock markets often arise from getting lucky on a single thousandfold gain, adding a lot of noise to the signal, but in prediction markets, prices are bounded between 0 and 1, limiting the impact of any one single event.

Better participants and better technology

Second, prediction markets themselves will improve. User interfaces have greatly improved already, and will continue to improve further. The complexity of the MakerDAO -> Foundry -> Catnip cycle will be abstracted away into a single transaction. Blockchain scaling technology will improve, reducing fees for participants (The [ZK-rollup Loopring](#) with a built-in AMM is already live on the Ethereum mainnet, and a prediction market could theoretically run on it).

Third, the demonstration that we saw of the prediction market working correctly will ease participants' fears. Users will see that the Augur oracle is capable of giving correct outputs even in very contentious situations (this time, there were two rounds of disputes, but the no side nevertheless cleanly won). People from outside the crypto space will see that the process works and be more inclined to participate. Perhaps even Nate Silver himself might get some DAI and use Augur, Omen, Polymarket and other markets to supplement his income in 2022 and beyond.

Fourth, prediction market tech itself could improve. [Here is a proposal from myself](#) on a market design that could make it more capital-efficient to simultaneously bet against many unlikely events, helping to prevent unlikely outcomes from getting irrationally high odds. Other ideas will surely spring up, and I look forward to seeing more experimentation in this direction.

Conclusion

This whole saga has proven to be an incredibly interesting direct trial-by-first test of prediction markets and how they collide

with the complexities of individual and social psychology. It shows a lot about how market efficiency actually works in practice, what are the limits of it and what could be done to improve it.

It has also been an excellent demonstration of the power of blockchains; in fact, it is one of the Ethereum applications that have provided to me the most concrete value. Blockchains are often criticized for being speculative toys and not doing anything meaningful except for self-referential games (tokens, with yield farming, whose returns are powered by... the launch of other tokens). There are certainly exceptions that the critics fail to recognize; I personally have benefited from ENS and even from using ETH for payments on several occasions where all credit card options failed. But over the last few months, it seems like we have seen a rapid burst in Ethereum applications being concretely useful for people and interacting with the real world, and prediction markets are a key example of this.

I expect prediction markets to become an increasingly important Ethereum application in the years to come. The 2020 election was only the beginning; I expect more interest in prediction markets going forward, not just for elections but for conditional predictions, decision-making and other applications as well. The amazing promises of what prediction markets could bring if they work mathematically optimally will, of course, continue to collide with the limits of human reality, and hopefully, over time, we will get a much clearer view of exactly where this new social technology can provide the most value.