

# REPLICATION CRISIS, PREDICTION MARKETS AND THE FUNDAMENTAL PRINCIPLE OF PROBABILITY

## OCTOBER 22, 2018

HARRY CRANE

**Goodhart's Law:** When a measure becomes a target it ceases to be a good measure.

- *Theory:* P-value measures the amount of evidence data provides for a scientific hypothesis.
- *Practice:* P-value < 0.05 is the target that determines whether results are published.

### Fundamental Principle of Probability

- (1) If you assign a probability to an outcome happening, then you must accept a bet on the other side at the correct implied odds.
- (2) Example: You claim that Clinton has 99% probability to win 2016 election (as Princeton Election Consortium did). This claim is meaningful only if you will offer 99-to-1 for a bet against Clinton.
- (3) Application to Replication Crisis: When publishing a result, authors provide:
  - (a) Replication criteria on which replication will be deemed successful.
  - (b) Replication probability  $p$  determines implied odds  $p/(1 - p)$  of bet against replication.
  - (c) Exposure Limit: Monetary amount that author backs up its claim.

### Prediction Markets

#### Will Bob Menendez be re-elected to the U.S. Senate in New Jersey in 2018?

Latest Price: 78¢ + 1¢



**Buy Yes** Click to match Offers starting at 76¢, or to make your own, lower Offer.

**Buy No** Click to match Offers starting at 22¢, or to make your own, lower Offer.

If this prediction comes true, PredictIt will redeem all Yes shares at \$1. Shares in No will have zero value. If this prediction does not come true, PredictIt will redeem all No shares at \$1. Shares in Yes will have zero value.

Data Rules Prices

Bob Menendez shall be the winner of the 2018 general election for U.S. Senator from New Jersey. PredictIt may determine how and when to settle the market based on all information available to PredictIt at the relevant time.

PredictIt reserves the right to wait for further official, party, judicial or other relevant announcements, reports or decisions to resolve any ambiguity or uncertainty before the market is settled. Markets may stay open or incur a delay in settlement well past the date of the contest in certain circumstances. If there is any change to an event, or any situation arises, that is not in PredictIt's view addressed adequately by the market rules, PredictIt will decide the fairest and most appropriate course of action.

PredictIt's decisions and determinations under this rule shall be at PredictIt's sole discretion and shall be final.

https://projects.fivethirtyeight.com/2018-midterm-election-forecast/senate/new-jersey/

ELECTION 2018  
**FiveThirtyEight** House forecast Senate Governor Midterms coverage More politics NEWS

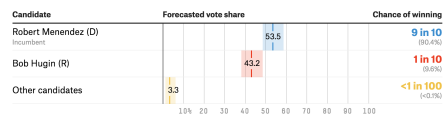
#### New Jersey LIKELY D

**9 in 10**

Chance the Democrat wins (90.4%)

**1 in 10**

Chance the Republican wins (9.6%)



**Prediction Market:**  $\Pr(\text{Menendez re-elected}) = 0.78$

**538 Forecast:**  $\Pr(\text{Menendez re-elected}) = 0.90$

**Which is more accurate?**

- If 538 estimate is right, the market is offering a 12% edge to bet on Menendez.
- Either (i) markets very inefficient or (ii) 538 estimates are unreliable. (Maybe both.)

**Applied to Replication Crisis:**

- (1) When a paper is submitted for publication, a betting market opens up.  
**Conditional on a replication attempt, the claimed results will replicate:**  
**YES: \$0.67**  
**NO: \$0.33**
- If replication attempt successful, **YES** worth \$1.00, **NO** worth \$0.00.
  - If replication unsuccessful, **YES** worth \$0.00, **NO** worth \$1.00.
  - If no replication attempt within (say) 1 year, then no action.
- (2) Rationale: The market incentivizes accurate assessment of the replication probability.  
(3) Journal uses market price in publication decision.  
(4) If accepted, market probability published along with the article.

**Comparison of FPP and Prediction Markets:**

	Prediction Market	FPP
Replication Probability	Market-driven	Set by authors
Role of Probability	Publication Decision, Signaling	Price of real bets Skin in the game
Consequence if Probability too high	Type I error: Publication Author benefits	Author loses money/funds Challenger gains
Consequence if Probability too low	Type II error: Rejection Author penalized	Conservative estimate No harm, no foul
Author incentive	Overstate probability Deception	Understate probability Conservative
Information asymmetry	Between Market & Author	Price in or risk ruin
Challenges to paradigm	Upper hand for "normal science" New ideas suppressed	Equal footing with <i>status quo</i>

**References:**

- H. Crane. The Fundamental Principle of Probability. <https://researchers.one/article/2018-08-16>
- H. Crane and R. Martin. In peer review we (don't) trust: How peer review's filtering poses a systemic risk to science. <https://researchers.one/article/2018-09-17>
- R. Hanson. Could Gambling Save Science?