Models & Markets

Comparing 2018 Midterm predictions

Question

- What is the best way to predict elections?
- When does each method perform best? Why?

Hypothesis

- Prediction **markets** will outperform **models**

Why Predict Elections?

1. Resource allocation
2. Strategy adjustment
3. Quantitative journalism
4. Uncertainty is scary!

How to Predict Elections?

1. Opinion polling
2. Poll aggregation
3. Forecasting models
4. Prediction markets

Markets Literature

- Focus on single race, mostly Presidential
- Compare to opinion polling

Room for Congressional comparison to model

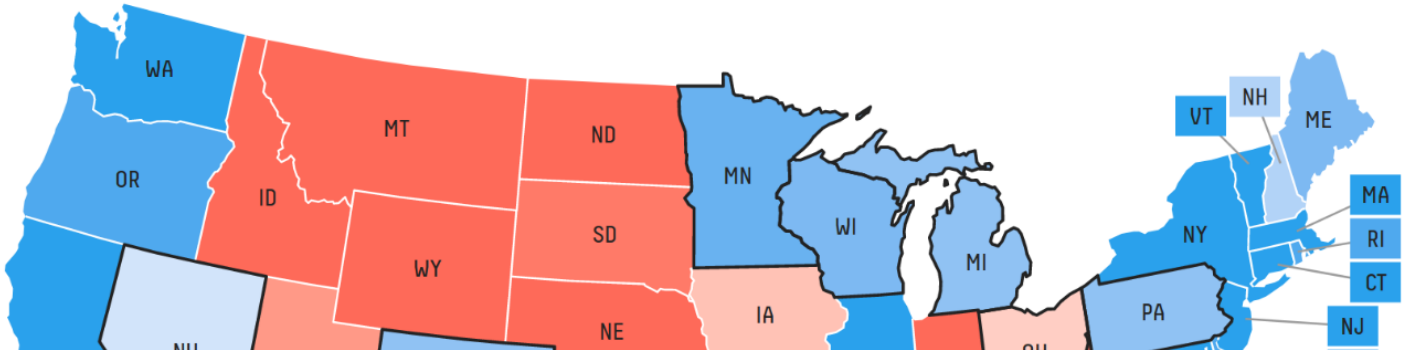
Forecasting Models

Forecasting Models

- 21st Century Invention
- Quantitative inputs
- Express uncertainty
- Probabilistic
- Proprietary

Who will win the presidency?

Chance of winning



FORECAST

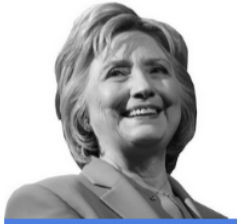
PRESIDENT SENATE

By Natalie Jackson and Adam Hooper

Additional design by Alissa Scheller

PUBLISHED MONDAY, OCT. 3, 2016 12:56 P.M. EDT

UPDATED TUESDAY, NOV. 8, 2016, 12:43 A.M. EST



CLINTON

98.0%

TRUMP

1.7%

Photos: Getty

Model Process

> *Most election models work in something like the following way: [1] they calculate the most likely outcome in a particular state... and then [2] they determine the degree of uncertainty around that estimate.*

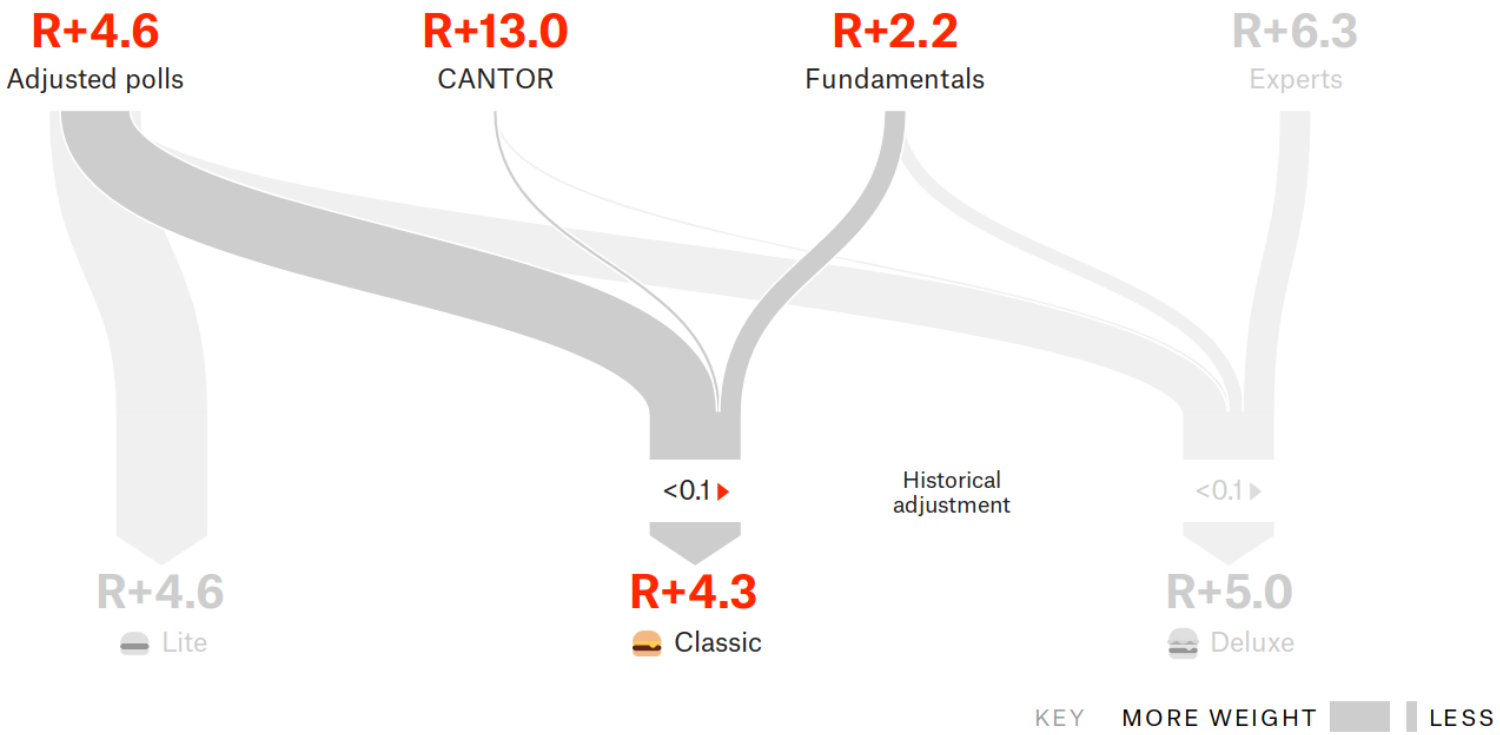
– Nate Silver

[1] Model Inputs

1. Weighted poll aggregation
2. CANTOR imputation
3. Fundamentals
    - incumbency
    - fundraising
    - previous election
    - scandals, etc

What goes into the Classic forecast in the New York 27th

The Classic version of our model projects a race's outcome by taking a weighted average of polls of a district (if available), polls of similar districts (CANTOR) and non-polling factors (fundamentals). It is then reverted toward a mean based on long-term trends in midterms and presidential approval ratings.



Model Variables

1. Date
2. Election
3. Party
4. Special
5. Incumbency
6. Voteshare(s)
7. **Probability**

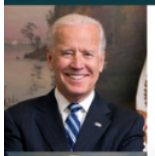
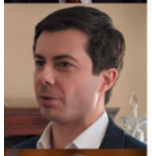
Prediction Markets

Prediction Markets

- Exchange-traded binary options
- Efficient market hypothesis
- Many purposes
  - 1503: Papal succession
  - 1988: Iowa Election Market
- Dubious legality
  - Academic waiver from CFTC

Market Contracts

- $850 limit
- Buyers on either side
- $1 or $0 based on outcome
- Sell at any time
- Price change

	Contract	Latest Yes Price	Best Offer		Best Offer
	Joe Biden	24¢ 1¢▲	25¢	Buy Yes Buy No	76¢
	Bernie Sanders	22¢ 1¢▼	22¢	Buy Yes Buy No	79¢
	Pete Buttigieg	18¢ NC	19¢	Buy Yes Buy No	82¢
	Kamala Harris	14¢ NC	15¢	Buy Yes Buy No	86¢
	Andrew Yang	9¢ 1¢▼	10¢	Buy Yes Buy No	91¢
	Elizabeth Warren	8¢ 1¢▲	9¢	Buy Yes Buy No	93¢

Market Variables

1. ID
2. Question
3. Name
4. Contract
5. Date
6. Volume
7. **Price**

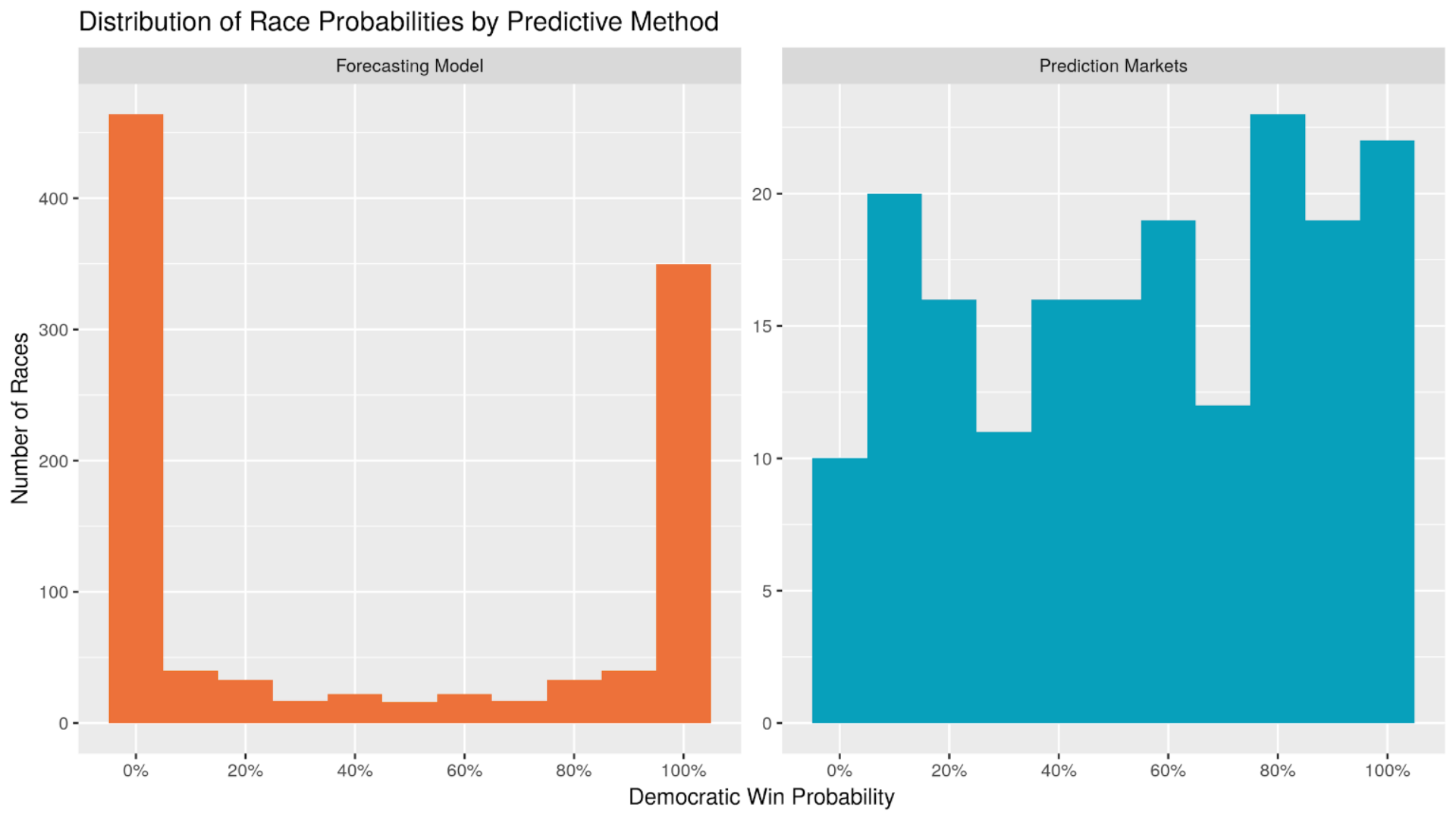
Comparing Methods

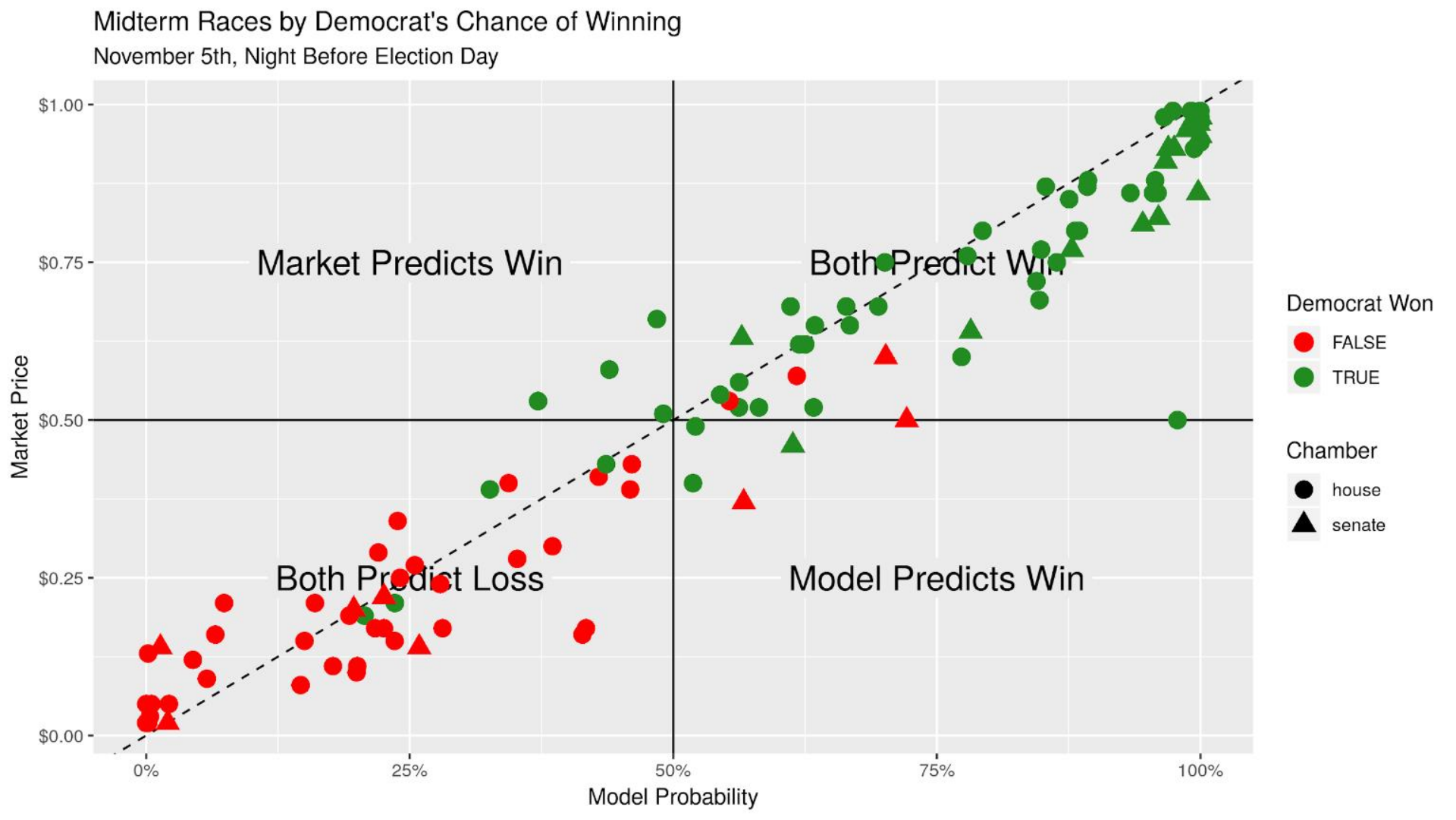
Models

1. Date
2. Election
3. Party
4. Voteshare(s)
5. **Probability**

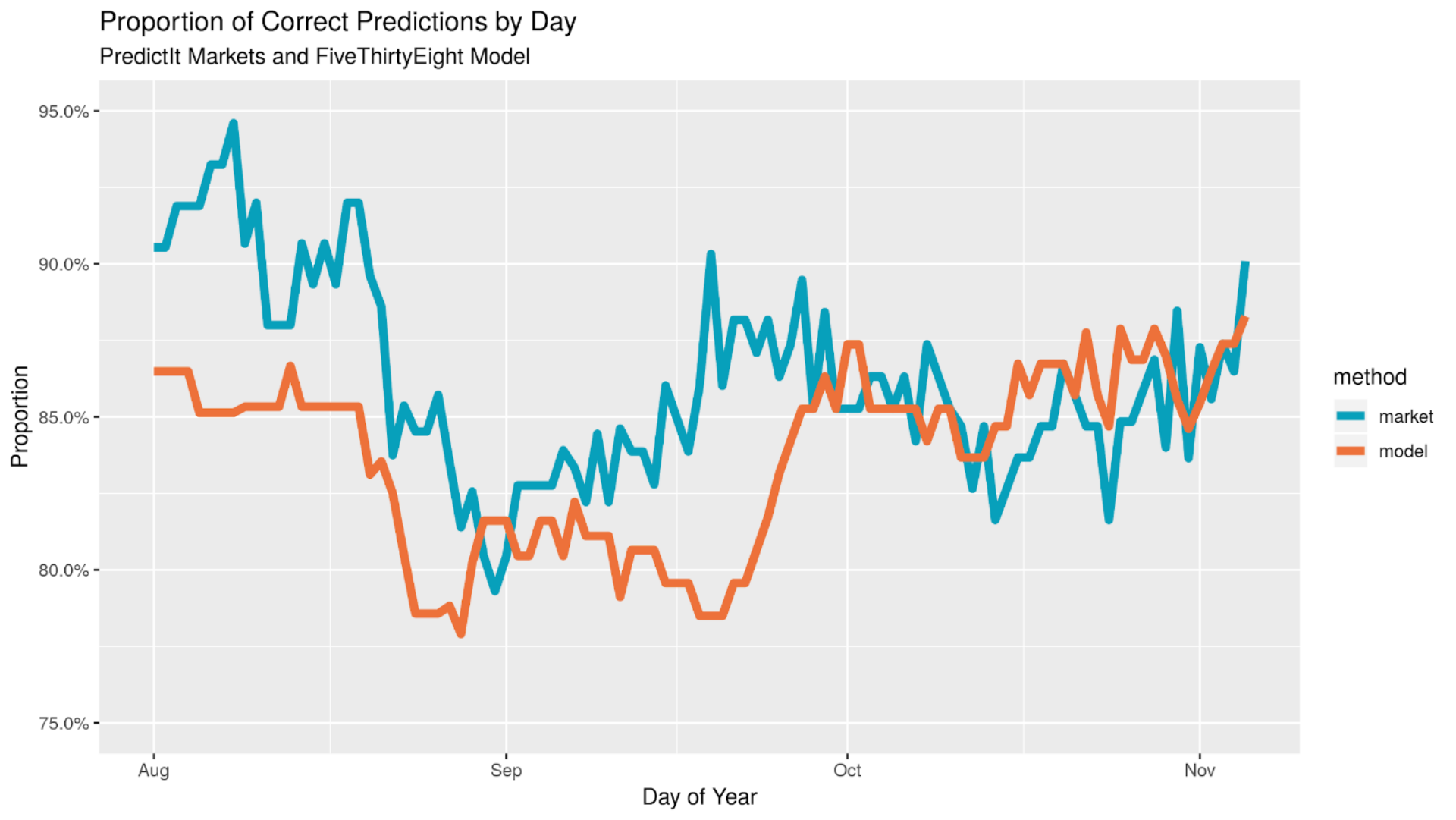
Markets

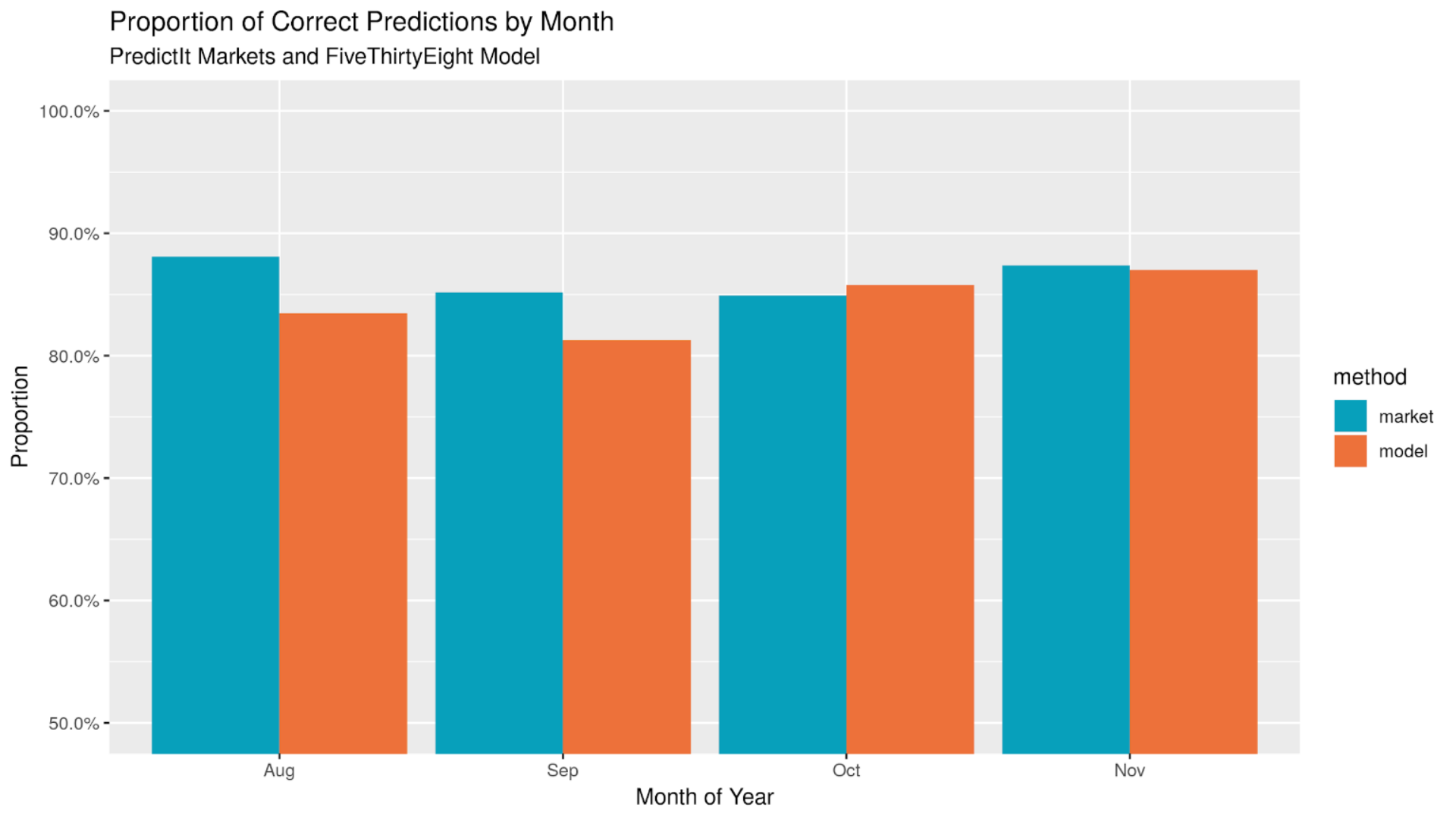
1. Date
2. Market
3. Contract
4. Volume
5. **Price**


Distribution of Race Probabilities by Predictive Method
Forecasting Model
Prediction Markets
Number of Races
0
100
200
300
400
0
5
10
15
20
0%
20%
40%
60%
80%
100%
Democratic Win Probability


Midterm Races by Democrat's Chance of Winning
November 5th, Night Before Election Day
Market Predicts Win
Both Predict Win
Both Predict Loss
Model Predicts Win
$1.00
$0.75
$0.50
$0.25
$0.00
Market Price
0%
25%
50%
75%
100%
Model Probability
Democrat Won
FALSE
TRUE
Chamber
house
senate

Date	Race	Method	Prob	Pred	Win	Hit
8/1	AZ-S1	Market	66	1	1	1
8/1	AZ-S1	Model	74	1	1	1
8/1	CA-12	Market	91	1	1	1
8/1	CA-12	Model	100	1	1	1
8/1	CA-22	Market	30	0	0	1
8/1	CA-22	Model	5	0	0	1
8/1	CA-39	Market	61	1	1	1
8/1	CA-39	Model	38	0	1	0


Proportion of Correct Predictions by Day
PredictIt Markets and FiveThirtyEight Model
95.0%
90.0%
85.0%
80.0%
75.0%
Proportion
Aug
Sep
Oct
Nov
Day of Year
method
market
model


Proportion of Correct Predictions by Month
PredictIt Markets and FiveThirtyEight Model
100.0%
90.0%
80.0%
70.0%
60.0%
50.0%
Proportion
Aug
Sep
Oct
Nov
Month of Year
method
market
model

Proportions Test

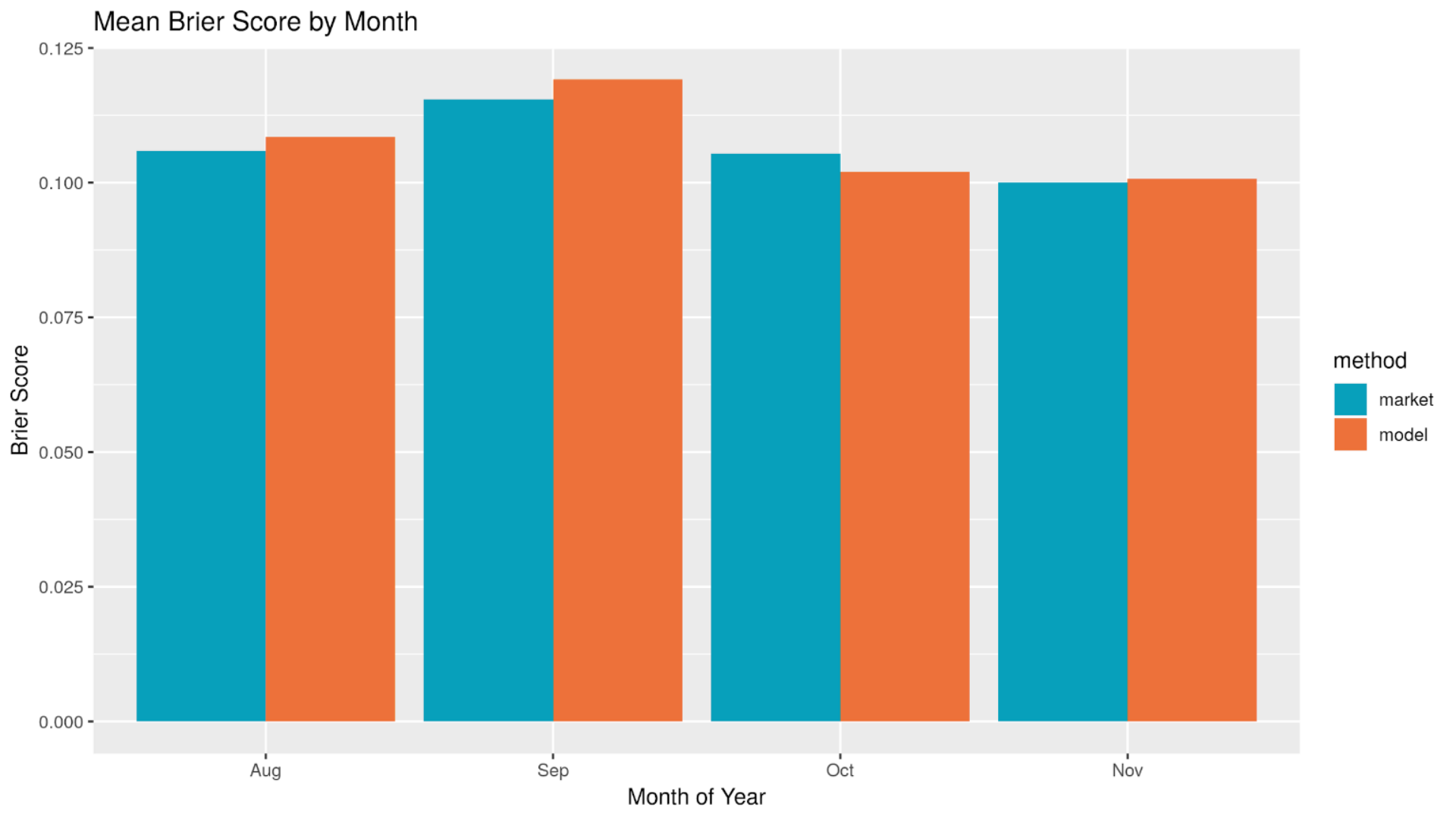
- Market proportion: 86.0343%
- Model proportion: 83.8057%
- Lower bound: 1.158%
- Upper bound: 3.2999%
- X-squared = 16.794
- p-value = 0.000042
- Alternative hypothesis: Two sided

	TRUE		FALSE	
	Markets	Model	Markets	Model
T	79.5 (51.2)	84.5 (51.2)	40.6 (9.7)	36.5 (9.7)
F	59.3 (4.3)	63.7 (6.5)	23.0 (34.3)	16.8 (32.1)

Brier Score

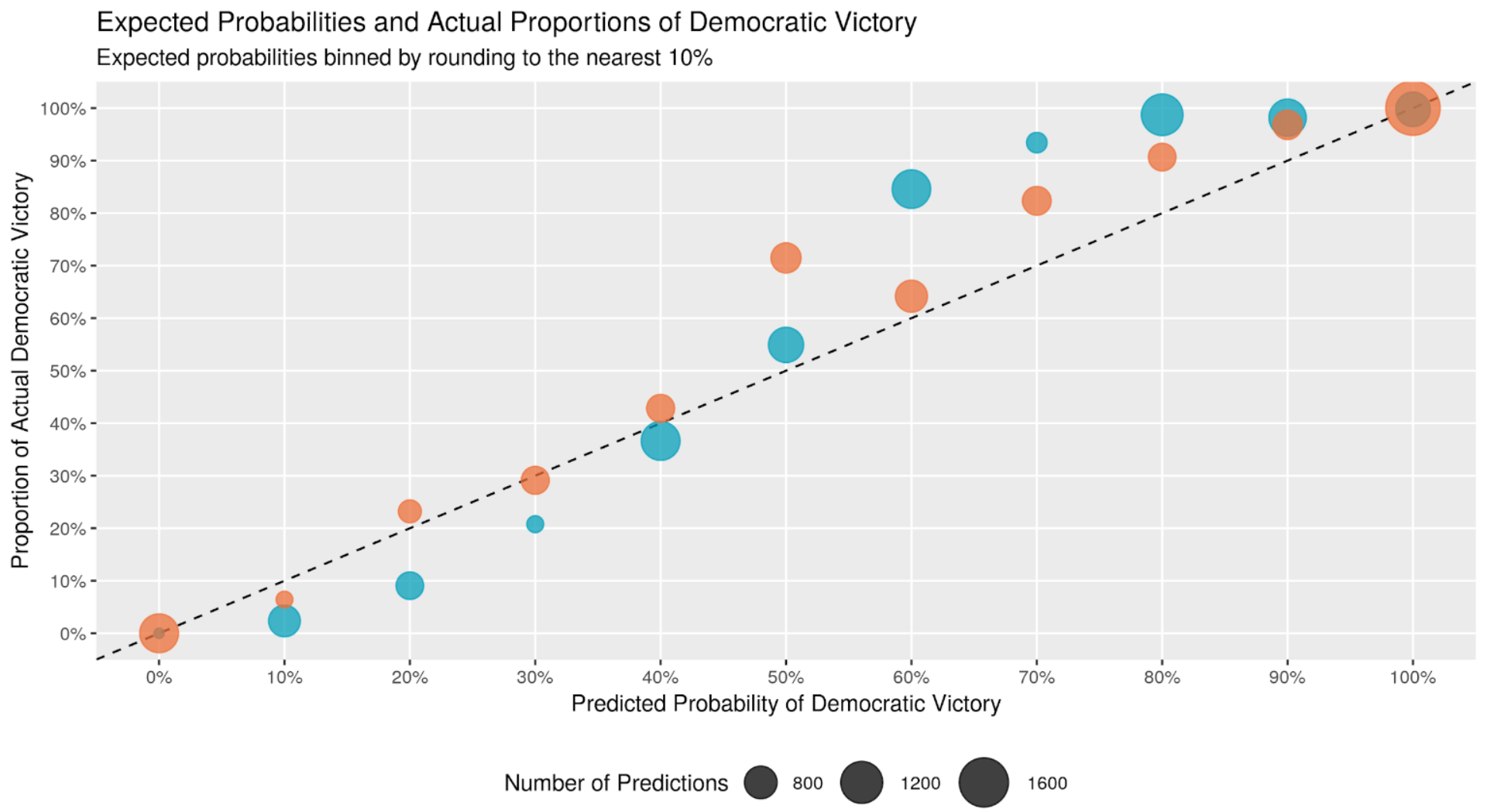
- Democrat has has 70% chance
- Democrat wins on election day
- Brier Score = $(0.70-1)^2 = 0.09$

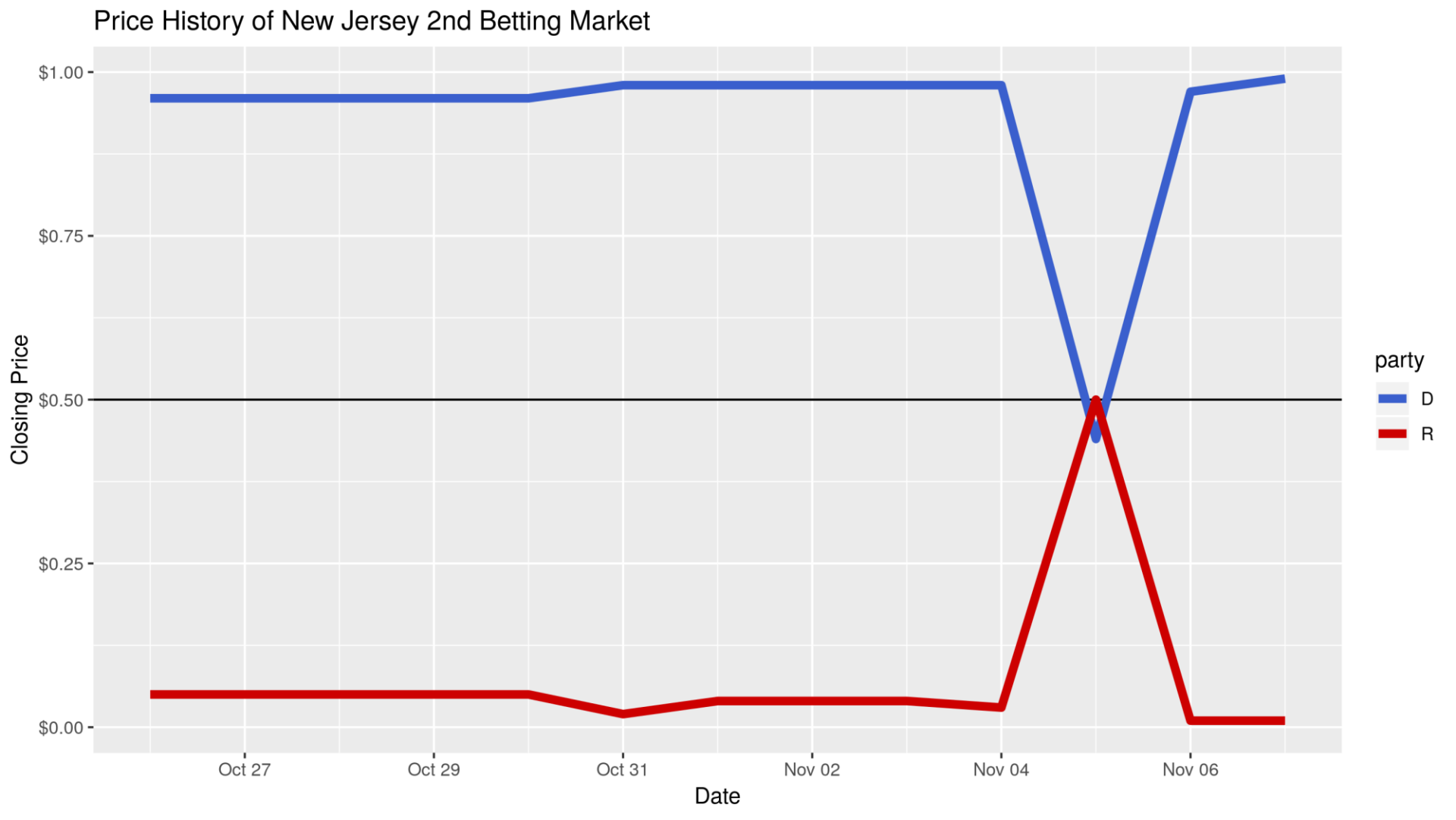
$$BS = \frac{1}{N}\sum_{t=1}^{N}(f_t - o_t)^2$$

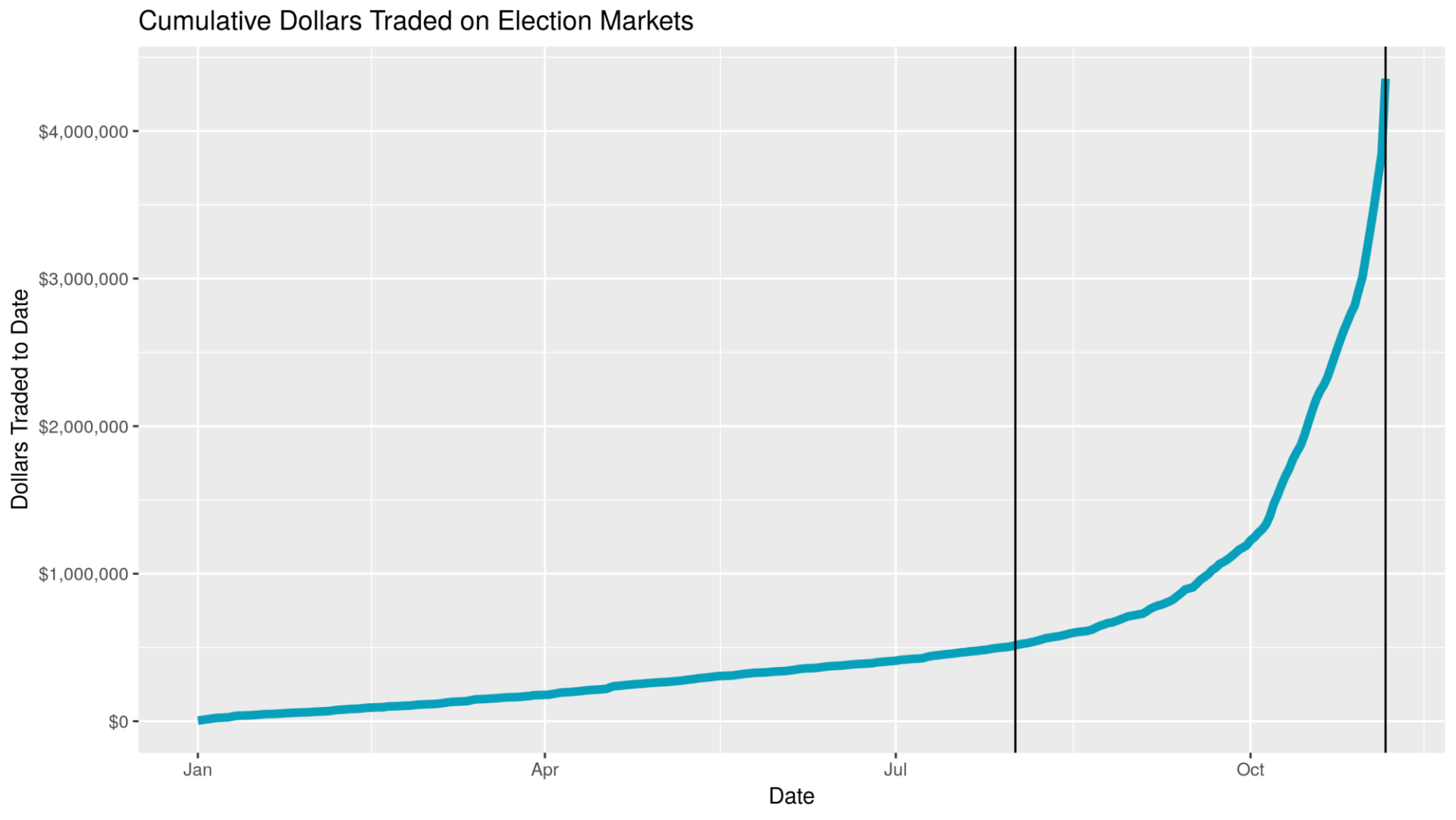

Mean Brier Score by Month
0.125
0.100
0.075
0.050
0.025
0.000
Brier Score
Aug
Sep
Oct
Nov
Month of Year
method
market
model

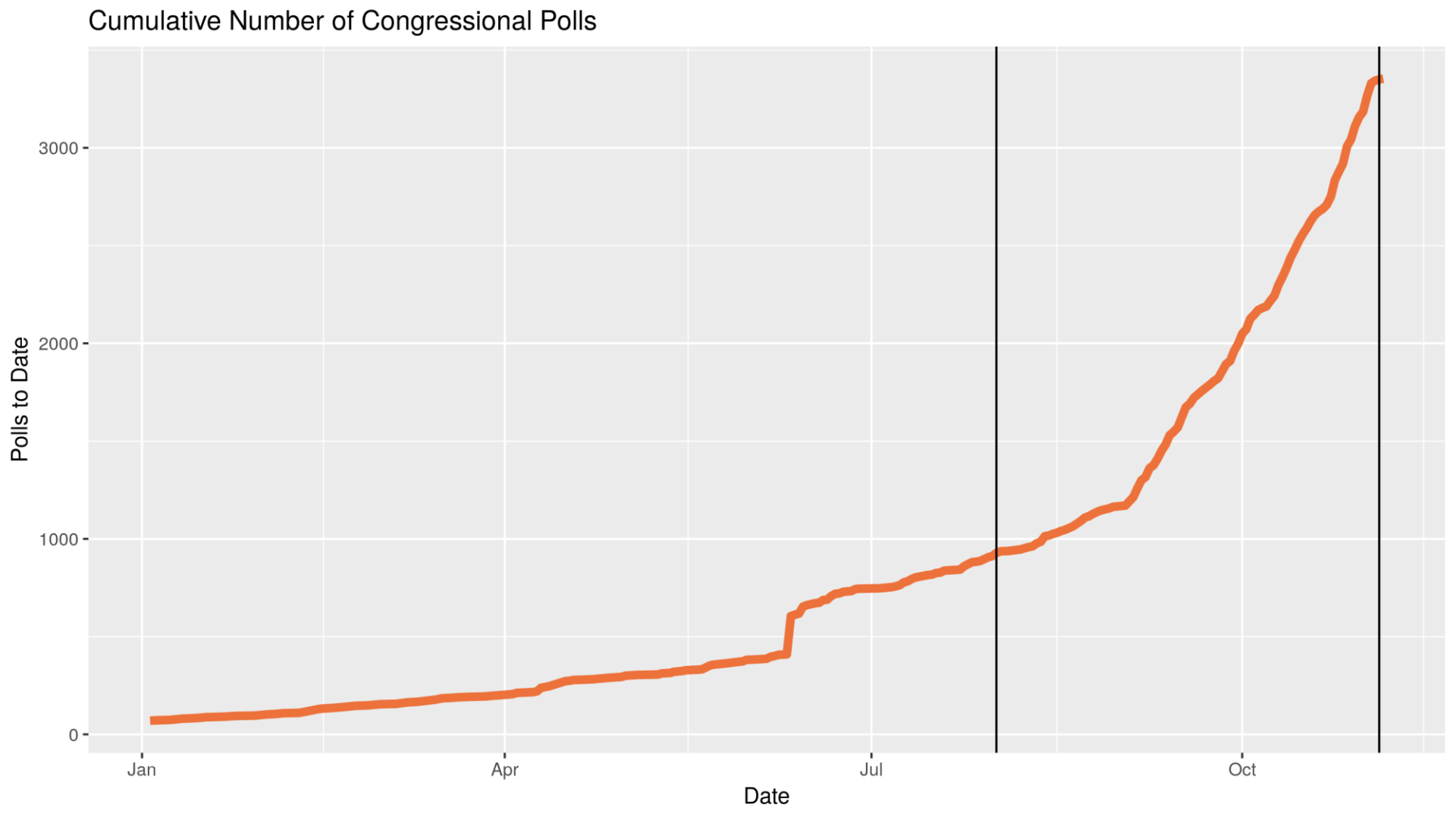
Brier T-Test

- Market mean: 0.1084
- Model mean: 0.1091
- Lower bound: −0.005
- Upper bound: 0.0035
- t = −0.33902
- p-value = 0.7346
- Alternative hypothesis: Difference


Expected Probabilities and Actual Proportions of Democratic Victory
Expected probabilities binned by rounding to the nearest 10%
Proportion of Actual Democratic Victory
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Predicted Probability of Democratic Victory
Number of Predictions
800
1200
1600


Price History of New Jersey 2nd Betting Market
$1.00
$0.75
$0.50
$0.25
$0.00
Closing Price
Oct 27
Oct 29
Oct 31
Nov 02
Nov 04
Nov 06
Date
party
D
R


Cumulative Dollars Traded on Election Markets
$4,000,000
$3,000,000
$2,000,000
$1,000,000
$0
Dollars Traded to Date
Jan
Apr
Jul
Oct
Date


Cumulative Number of Congressional Polls
Polls to Date
3000
2000
1000
0
Jan
Apr
Jul
Oct
Date