Data Science Quarterly: Manuscript for Research Section
Volume 1, Issue 1, January 2021
ISSN 2693-8871 (print) ISSN 2693-8863 (online)

Predicting the 2020 Presidential Election

Thomas W. Miller

Working paper revised 24 November 2020

Abstract This paper reviews three election forecasting methods: economic and voter behavior models, opinion polls, and prediction markets. It cites literature relating to election forecasting with a focus on prediction markets. And it shows how to use pricing data from prediction markets to forecast the winning ticket in the U.S. presidential election of 2020.

Keywords Prediction markets · Election forecasting · Electoral College · Presidential election · Monte Carlo simulation

1 Introduction

What methods should we use to forecast winners in the last few months leading up to an election? There are three general approaches to consider: (1) models that draw on historical trends in demographics, economics, and voting behavior, (2) opinion polls, and (3) prediction markets. We argue that prediction markets hold the most promise.

What about predictions for the 2020 presidential election? Few have doubts about the direction of the nationwide popular vote. Most experts believe that the Democratic ticket will draw more popular votes than the Republican ticket. Furthermore, the percentage advantage of the Democratic ticket is of interest only to a limited number of people: candidates for Congress, statewide, and local elections (due to "coattail" effects), as well as political scientists, modelers, and traders in vote-share prediction markets. What is in doubt is which ticket will win in the Electoral College. We propose a straightforward method for picking the winning ticket in the 2020 U.S. presidential election, a method that builds on Electoral College prediction markets and statistical simulation.

The author thanks representatives of PredictIt (www.predictit.org) for providing market data relating to this research.

2 Literature Review

Modelers rely on historical data relating to demographics, economics, and voting behavior. Modelers rely on secondary research, and they believe that accurate forecasts can be obtained well in advance of political campaigns. Pollsters conduct primary research, trying to gauge public opinion about candidates. There are also hybrid models that rely on historical observations and current polls. We provide a general review of the work of modelers and pollsters.

Prediction markets provide an alternative to models and polls. Making a forecast from prediction markets involves collecting information about current betting prices. No adjustments are required to conform to underlying theory, past history, or opinion polls. We review the history of prediction markets, efficient markets theory, and empirical results, setting the stage for a new approach to forecasting the U.S. presidential election.

2.1 Modelers

Political scientists point to observed relationships between demographics and voting Democrat or Republican. A voter's race, income level, age, and sex are related to party affiliation and candidate preferences. Economic conditions affect elections, with periods of growth in gross domestic product associated with incumbent party victories over challengers. There is a rich literature demonstrating election forecasting models (Abramowitz 2008; Campbell 2014; Fair 1978, 1996, 2011; Holbrook 2008, 2012; Jerôme and Jerôme-Speziari 2012; Rosenstone 1983).

There is extensive research to support voting behavior trends that may be incorporated into forecasting models. Theiss-Morse et al. (2018) documented trends in political attitudes and behavior using a comprehensive longitudinal study, the American National Election Studies (ANES). Surveys were carried out during every national election year since 1952, with the exception of 2006, 2010, and 2014. The data are publicly available from the Interuniversity Consortium for Political and Social Research (ICPSR 2020). Selected findings from ANES, documented by Theiss-Morse et al. (2018), include the following:

- trust in government has declined substantially, especially among young adults;
- women have increased their participation in elections and, since 2008, have participated in higher percentages than men;
- voter turnout among Whites has been fairly constant, whereas voter turnout among Blacks has increased substantially;
- across the entire period of study, voter turnout for persons under 35 is substantially lower than voter turnout among persons 35 and older;
- the percentage of adults identifying as Democrats has declined;

- the percentage of adults identifying as Independents has increased, so that the percentage of Independents now exceeds the percentage of Democrats and the percentage of Republicans;
- for persons residing in the South, the percentage of adults identifying as Republican has increased; and
- since 1968, the first year that there was a survey item regarding positive/negative views of the presidential candidates, the presidential election of 2016 was the only election in which adults had, on average, a negative view of both the Democratic and Republican candidates.

Lewis-Beck and Tien (2008) note that most election forecasting models incorporate economic factors and prior voting patterns. Models commonly include economic explanatory variables such as changes in gross national product, inflation rate, various leading economic indicators, and personal income growth (a voter pocketbook variable). Many models include explanatory variables relating to political parties and recent voting results, such as the presence of an incumbent, the number of terms the incumbent party has been in office, primary election results, and the popularity of candidates.

Allan J. Lichtman (2008, 2020) presents an election model that relies on thirteen binary predictors. The predictors, called "keys to the White House," relate to general economic, political party, and candidate characteristics—all factors that can be observed well in advance of an election. Lichtman's model, by his own admission, is retrospective, an attempt to explain the past more than forecast the future. Nonetheless, his model correctly predicted the popular vote winner of every election since 1860. Polling data are excluded from Lichtman's model, as are campaign activities. If we are to believe Lichtman's approach, we would have to discount political rallies and party platforms, as well as candidate speeches and debates, as affecting election outcomes. In today's electoral environment, with Democratic candidates for president assured of popular vote victories, Lichtman (2020) recognizes the need to predict the winner of the Electoral College vote.

Gelman and King (1993) noted that the popular vote outcome of a U.S. presidential election can be predicted within a few percentage points using information available months before the election. Predicting the nationwide popular vote is straightforward compared to predicting Electoral College results.

Historical trends and relationships are of little value in predicting what will happen in the last few months of a campaign, and what happens in the last few months of a campaign can affect Electoral College results. In many states (the"swing states"), voter registrations are divided more evenly between Democrats and Republicans, and in most states, there are large numbers of Independents. Small within-swing-state changes in voting behavior may well determine the outcome of the U.S. presidential election.

FiveThirtyEight, a data journalism organization and website led by Nate Silver, builds forecasting models for various events, most notably in politics

and sports. For the U.S. presidential election of 2008, FiveThirtyEight gained recognition for the accuracy of its forecasts.

Silver's recommendations for modelers are documented in *The Signal and the Noise* (Silver 2012): think probabilistically (preferably as a Bayesian), update forecasts as new data become available, and look for consensus (by combining information from many sources). Silver (2012, 67) describes a hybrid election forecasting approach that combines polling data with information about the economy, demographics, and voting patterns of states. FiveThirtyEight's (2020b) models utilize statistical simulations with forecasts for each Electoral College market.

Another hybrid modeling approach is *The Economist's* presidential forecast (*Economist*, Gelman, and Heidemanns 2020). This model incorporates national and statewide voting patterns from the past, as well as up-to-date national and statewide polls (Gelman et al. 2020).

Miller (2008) identifies providers of research and information services along two dimensions: primary versus secondary research and theory-driven versus data-driven research. Traditional modelers rely on secondary research, whereas pollsters conduct primary research. Traditional modelers and pollsters, coming from history and the social sciences, tend to be theory-driven more than data-driven. Researchers associated with FiveThirtyEight and *The Economist's* presidential forecast typify a new breed of modeler, more data-driven than theory-driven. This new breed of modeler may call on theory in specifying Bayesian priors, but most of the modeler's time is spent updating priors with recent data from opinion polls.

2.2 Pollsters

Pollsters hope that respondents (or weighted groups of respondents) comprise a representative sample of likely voters. Political scientists often complain about opinion polls, as do politicians and analysts. Opinion polls try to predict the behavior of likely voters on election day. Polling has undergone substantial changes over the years, as documented in the political science and survey research literature (Jacobs and Burns, 2004; Jacobs and Shapiro 1995, 2005; Hillygus 2011; Traugott 2014).

Some pollsters work for political candidates and parties. Others are independent organizations, often using polls as a way to promote business services, including marketing research and management consulting. Still others are aligned with universities or media organizations.

John F. Kennedy's presidential campaign of 1960 was the first to use political polling as a planning tool (Honomichl 1990, 48). Kennedy employed Louis Harris as his main pollster. From the 1960s through 1980s, Peter D. Hart and Patrick H. Caddell were prominent pollsters working on behalf of Democratic candidates. Their counterparts on the Republican side were Peter M. Teeter and Richard B. Wirthlin (Honomichl 1990).

George H. Gallup Sr. founded the Gallup Organization in 1935 and published the first Gallup Poll in that same year (Honomichl 1990). Insisting on politically balanced reporting, Gallup refused to conduct surveys commissioned by Democratic or Republican parties. The Gallup Poll established a reputation for accurate election forecasting through the use of statistical sampling and survey best practices. In 1936, Gallup correctly predicted the election of Franklin Roosevelt over Alfred Landon, in contrast to the incorrect and now infamous forecast of the *Literary Digest* (Crossley 1937).

The Gallup Poll long held a prime position among pollsters, with telephone interviewing as its primary modality. The Gallup Poll achieved wide distribution of its election forecasts across various media outlets nationwide from 1936 onward. In 2012, the Gallup Poll incorrectly predicted that Mitt Romney would defeat Barack Obama. Prior to the 2016 presidential election, the Gallup organization announced that it would no longer conduct horse-race polling (White 2015), but Gallup continues to conduct surveys relating to voter opinions.

Today's political polling landscape is characterized by FiveThirtyEight (2020a), with pollsters identified by their use of one or more of the following data collection modalities:

- (Live) live telephone interviews (perhaps with cell phones);
- (Landline) live telephone interviews, not including cell phones;
- (IVR) interactive voice response, otherwise known as automated polls or "robopolls";
- (Online) polls conducted through the Internet; generally, this means by a web browser (inclusive of text messaging and application-based polling on mobile phones); and
- (Mail) by the United States Postal Service or other physical mail service.

FiveThirtyEight (2020a) assigns letter grades to pollsters based on their use of proper survey technologies and their history of accurate election forecasts. The organization also assesses each pollster in terms of its bias in favor of Democratic or Republican candidates. As of May 19, 2020, the organization maintained a list of 453 pollsters, with ratings based on 9,559 polls.

Well-designed statistical surveys are essential to the proper operation of political polls. Accuracy can be affected by sampling and nonsampling errors, including coverage, nonresponse, measurement errors associated with respondent and interviewers, and postsurvey processing errors (Weisberg 2005; Groves et al. 2009). Modality differences are a concern as well because telephone and online surveys have the potential of yielding different results (Miller 2001; Miller and Dickson 2001).

Unfortunately, polls are costly and inefficient. Even with newer data collection modalities, opinion polls can take days or weeks to complete. There is the initial planning of the survey instrument and sampling scheme, followed by data collection, which may include multiple attempts to reach individual respondents. The analysis can also be time-consuming because there may be

a need to weight survey responses to ensure that reported results are representative of the voting population. It is not surprising, then, that polls are out-of-date as soon as they are published.

There is wide variability in polling results, partly due to different methodologies, partly due to statistical variability, and, as Gelman and King (1993) note, partly due to what voters learn through the media. Polls are most useful when conducted just prior to an election. Averages across polls conducted just prior to an election usually provide accurate election forecasts of nationwide popular vote percentages.

Many think the 2016 U.S. presidential election was a failure of opinion polling. Not so. Popular vote forecasts for the 2016 election were highly accurate, with national poll averages anticipating the final election results: Hillary Clinton received 48.0 percent of the popular vote and Donald Trump 45.9 percent (Theiss-Morse et al. 2018). To pick the winner of a U.S. presidential election, however, it is not sufficient to anticipate the nationwide popular vote. The winning ticket is determined by the Electoral College.

2.3 Prediction Markets (Overview)

Prediction markets go by various names: artificial markets, decision markets, electronic markets, event markets or event futures, forecasting markets, information markets, betting markets, and virtual stock markets. Prediction markets provide contracts linked to future events.

There is extensive literature discussing these markets, including articles in many academic journals and specialized journals such as *The Journal of Prediction Markets* and *Electronic Markets—The International Journal on Networked Business.* Tziralis and Tatsiopoulos (2007) identified 155 articles relating to prediction markets between 1990 and 2006, and Horn et al. (2014) identified an additional 304 articles relating to prediction markets between 2007 and 2013.

Rhode and Strumpf (2004) trace the history of betting on U.S. presidential elections. Between the Civil War and World War II, political betting exchanges, although often illegal, were active and in the public eye. Newspapers would provide price quotations in the final month of each campaign. Betting pools were often in excess of tens of millions of dollars.

The most extreme example of political betting was the U.S. presidential election of 1916. That election pitted Supreme Court Justice and former Governor of New York, Charles Evans Hughes, the Republican candidate, against New Jersey governor, former Princeton professor, and the eventual Democratic winner, Woodrow Wilson. Records show that more than $237 million (2020 dollars) was bet on the 1916 election, more than twice the combined spending by the candidates' campaigns. Betting on that election exceeded transactions on stock exchanges during the same period (Rhode and Strumpf 2004).

The Iowa Electronic Markets, a collection of nonprofit prediction markets at the University of Iowa, began as an experiment in 1988 with permission from

the Commodity Futures Trading Commission (CFTC). During its first year of operation as a real-money prediction market, it was open only to University of Iowa students, faculty, and staff. Subsequently it opened to adult traders worldwide. The Iowa Electronic Markets offers winner-take-all and vote-share contracts, as well as conditional contracts defined across sequences of events, all with a betting limit of $500. While the Iowa Electronic Markets platform offers general, nationwide contracts for the U.S. presidential election, there are no contracts for individual Electoral College markets.

PredictIt is an experimental project operated for academic purposes, also with permission from the CFTC. PredictIt's home is Victoria University of Wellington, New Zealand. PredictIt, which describes itself as "the stock market for politics," offers more than three hundred political futures contracts. Each contract concerns a specific political event and bears an individual share price between 1 and 99 cents. Share prices change over time, reflecting traders' beliefs about the probabilities of political events. Investments in each contract is limited to $850.

Most relevant to the U.S. presidential election are PredictIt contracts for the fifty-six Electoral College markets. For each electoral market, a trader can purchase shares for the Democratic ticket or the Republican ticket. After the official election results are declared for a market, PredictIt will close the contract for that market, at which time a share for the winning ticket is worth $1 and a share for the losing ticket is worth nothing. PredictIt charges a 10 percent fee on profits and a 5 percent fee on withdrawals.

The Iowa Electronic Markets and PredictIt support research about prediction markets by providing free access to market data for academic and non-profit organizations. Many economists argue that prediction markets should be more widely available because they provide a plentiful source of information to guide public policy (Abramowitz 2004; Hahn and Tetlock 2006; Arrow et al. 2008).

Betfair is a for-profit online gambling company with headquarters in London, UK. It offers futures contracts for sporting and political events and charges various commissions and fees for its services. Included among the political events are contracts for congressional and presidential elections. Betfair offers winner-take-all contracts for the fifty states of the United States rather than for the full set of fifty-six Electoral College markets.

There are additional exchanges to consider for future elections. FTX Exchange, a financial derivatives and cryptocurrency futures exchange available outside the United States, offers political futures contracts, including contracts on who will win the U.S. presidential election. There are also emerging decentralized exchanges built on blockchain protocols: Polymarket, Omen, Gnosis, and Augur. None of these exchanges offers trading in Electoral College markets, as needed for forecasting the outcome of the U.S. presidential election.

2.4 Prediction Markets (Efficient Markets Hypothesis)

The efficient markets hypothesis is often used to explain why prediction markets provide accurate predictions of future events. The efficient markets hypothesis, which draws in part from Hayek (1945), connotes that market prices reflect all information known to traders.

Prices in an efficient market, as defined by financial theory (Fama 1970, 1991), will be good predictors of future events. To ensure efficiency, a prediction market should be open to a large and diverse group of traders and resistant to manipulation by individual traders. Information should be plentiful and widely shared. An efficient market, by definition, incorporates all available information about future values. Indeed, we can argue that the purpose of prediction markets is to aggregate information, producing prices that can be used to forecast future events.

We can argue that prediction markets are better than polls because prediction markets can respond to all factors relevant to an election. While opinion polls try to anticipate the behavior of likely voters on election day, prediction markets consider voting practices and contingencies as well as voter intentions. Prediction markets can also reflect voter demographics and economic conditions.

Well-designed prediction markets provide relevant, up-to-date information, responding to all events in the public sphere. If a candidate gives an especially good speech or convention events play well in the media, prediction market prices respond. If a candidate misspeaks or says something that reflects poorly on his or her character, that, too, can move prediction market prices. Weather forecasts for election day could affect prediction market prices, as could election regulations, election meddling, manipulation of voter registration rolls, vote-by-mail restrictions or delays, and reductions in voting locations or voting hours. Anything known to the public can affect prediction market prices. That includes what traders learn from the media, what they hear from pollsters, analysts, and pundits.

A believer in efficient markets might say that political prediction markets reflect collective information across all people willing to place bets on the outcomes of political contests. Prediction markets reflect the "wisdom of the crowd," a concept introduced by Surowiecki (2005). The wisdom of the crowd, revealed through participation in markets, is superior to deliberation by groups of analysts, experts, or pundits. In fact, groups of like-minded individuals are susceptible to "group think," with members encouraging one another to accept extreme opinions (Sunstein 2005, 2006a, 2006b).

The efficient markets hypothesis has its detractors (Manski 2006) and its supporters (Wolfers and Zitzewitz 2004, 2006b). To say that prediction markets are efficient is not to say that all traders are efficient or rational. They are not. It is sufficient for traders at the margin to be informed and rational (Olivea and Rietz 2004; Wolfers and Zitzewitz 2006a). Information efficiency in betting markets has been an active area of research (Williams 2005; Blackwell and Pickford 2011).

2.5 Prediction Markets (Empirical Results)

We do not have to believe that prediction markets are efficient to use prediction market prices in making election forecasts. There is ample evidence that prices from political prediction markets provide accurate election forecasts.

The historical record of fifteen presidential elections from 1884 to 1940 shows that in eleven of these elections the favored candidate won. The underdog won only once, in the 1916 Hughes-Wilson contest. And odds were essentially even in the remaining three contests (Rhode and Strumpf 2004).

Equally impressive has been the predictive performance of the Iowa Electronic Markets (IEM), as documented by Forsythe et al (1992), Berg and Rietz (2006), Berg et al. (2008), and Stix (2008). Berg and Rietz (2006) identified "stylized facts" from the IEM:

- compared with the general population, traders are more highly educated, from higher income levels, and more likely to be male than female;
- trader biases/preferences affect portfolio choices;
- while as many as twenty percent of IEM trades could have been executed by robots, there is no reason to believe that robots introduce pricing bias;
- there is no strong evidence of price manipulation in the IEM;
- IEM prices respond quickly to news; and
- as predictors of election outcomes, IEM prices often outperform political polls.

Berg et al. (2008) studied the five presidential elections from 1988 through 2004, comparing vote forecasts from the Iowa Electronic Markets with vote forecasts from 964 political polls. They found that the prediction markets provided more accurate forecasts 74 percent of the time. Furthermore, prediction markets were substantially better than polls more than one hundred days prior to elections.

Many economists and social researchers have observed that predictive markets outperform opinion polls in picking winning candidates (Kou and Sobel 2004; Leigh and Wolfers 2006; Arrow et al. 2008; Luckner et al. 2012; Rothschild 2009; Arnesen and Bergfjord 2014). There is evidence, as well, that prediction markets outperform models that rely on polling data (Crane:2018).

Even critics of prediction-market-based forecasts show that prediction market prices and opinion poll results just prior to an election have very high correlations (Erikson and Wlezien 2008a, 2012). Prediction markets provide an efficient and accurate approach to election forecasting. Figure 1 provides a summary of what we know about opinion polls versus prediction markets.

Fig. 1 Opinion Polls versus Prediction Markets

Opinion Polls	Prediction Markets
Representative samples of likely voters	Any adult worldwide who wants to bet (there can be programmatic trading from software robots) human traders tend to be better educated, from higher income levels, more likely male than female
Respondent volunteers, no compensation	Traders bet with money, hoping to make a profit after fees are collected
Respondents are asked to reveal their opinions about issues and candidate preferences, how they would vote if the election were held today	Markets operate online
Various modalities employed, including phone, voice response, mail, and online	Trading is available continuously and market prices vary continuously
Can take days to complete, beginning with survey and sampling design, data collection, and data analysis (may involve respondent weighting to ensure representative samples)	Predictions can be updated instantly in response to campaign events
Poor predictions early in campaigns	Good predictions early in campaigns
Just prior to the election, averages across polls can provide good predictions about election outcomes	Excellent predictions late in campaigns

3 Methods

Two things are required to pick the winning ticket in a U.S. presidential election: accurate forecasts of popular winners vote in each electoral market and simulation modeling. We estimate the probability of a Democratic and Republican victory in each Electoral College market and then use those probabilities in a statistical simulation.

PredictIt provides prices for the Democratic and Republican tickets in each of the 56 Electoral College markets: 48 of the 50 states, three regions in Maine, four regions in Nebraska, and the District of Columbia. We convert prediction market prices for the Democratic and Republican tickets into estimated probabilities of winning. Let D be the latest price for the Democratic ticket and let R be the latest price for the Republican ticket. Then the estimated probability of a Democratic victory is $D/(D+R)$ and the estimated probability of a Republican victory is $R/(D+R)$.[1]

[1] Wolfers and Zitzewitz (2006b) provide mathematical justification for the assertion that prediction market prices correspond to mean beliefs among traders. This theoretical result applies across a variety of utility functions. Appendix 1 shows presidential election data from the 56 Electoral College markets, with *state* identifying the market, *votes* the number of Electoral College votes, *dem* the latest price for the Democratic ticket, *rep* the latest price for the Republican ticket, and *dempro* and *repprob* being estimated probabilities for the Democratic and Republican tickets, respectively. Data in the appendix represent prediction market prices on the morning after the Republican National Convention, August 28, 2020.

Table 1 Forecast Verification (Brier Scores)

Forecasting Method	Brier Score
Naïve 50-50 Forecast	0.250
2016 Election Repeat	0.172
PredictIt-Based Estimates	0.041
FiveThirtyEight Final Model	0.040

Data sources, November 3, 2020: https://www.predictit.org
https://projects.fivethirtyeight.com/2020-election-forecast/

With individual state or market forecasts expressed as probabilities, we can use the Brier score, a quadratic scoring metric, to assess their quality (Brier 1950; Selten 1998). The Brier score is a strictly proper scoring metric used in assessing the quality of probability forecasts, such as those employed in weather forecasting (Murphy 1993; Gneiting and Raftery 2007; Casati et al. 2008). Let p_i be the forecasted probability of a Democratic victory in a market, and let δ_i be the dichotomous outcome variable, where $\delta_i = 1$ if the Democratic ticket wins the election and $\delta_i = 0$ if the Democratic ticket loses the election. The resulting Brier score varies from 0 to 1, with 0 being the best possible score. The Brier score is typically computed over a set of n bins for probabilities:

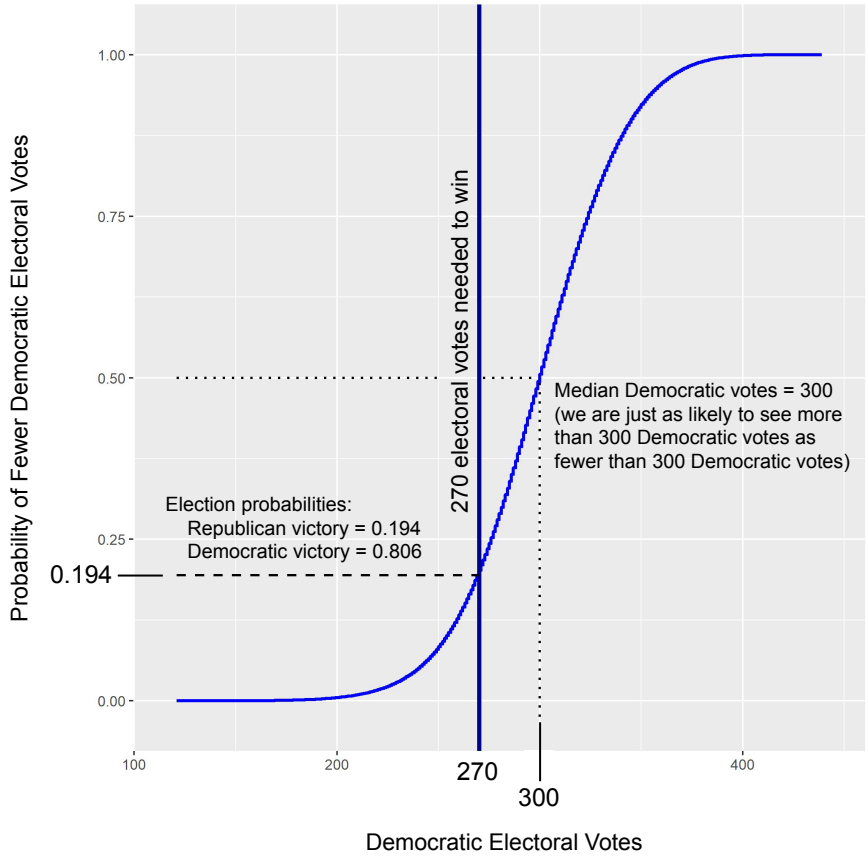
$$S = \sum_{j=1}^{n} (p_i - \delta_i)^2$$

With each update of prediction market prices and associated probabilities, we can use statistical simulation to generate one million hypothetical elections. We note the total electoral votes for the Democratic and Republican tickets across these elections. Our simulation methods are consistent with methods employed by other election forecasters. What is different about our method is a reliance on prediction markets rather than opinion polls.

4 Results

Table 1 shows Brier forecast verification results for four probabilty forecasts across the 56 electoral markets. An initial baseline forecast represents a naïve 50-50 forecast, which is equivalent to flipping a fair coin for every electoral market. A second baseline forecast sets each electoral market probability to be the observed relative frequency of Democratic versus Republican votes from the 2016 presidential election. The third probability forecast is obtained from PredictIt prices just prior to the opening of the polls in the east on the last day of voting, November 3, 2020. And the fourth is from the final FiveThirtyEight model on that same day. A straightforward prediction-market-based model is able to provide probability forecasts that are comparable to the forecasts of FiveThirtyEight, a complex opinion-poll-driven model.

Fig. 2 Results from One Million Hypothetical Elections

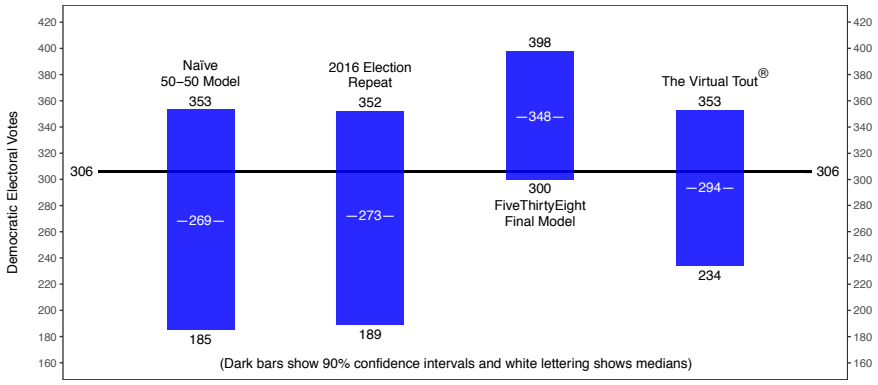


Data source: https://www.predictit.org, August 28, 2020

Forecast verification a la Brier treats all probabilties equally. But electoral market probabilities are not of equal value because electoral votes vary from one market to the next. A meaningful comparison of presidential election forecasting methods requires statistical simulation. We must run hyoptheitcal elections to see which method is most accurate.

Figure 2 summarizes results from one million hypothetical presidential elections based on prediction market data from the morning after the Republican National Convention, August 28, 2020. Because 270 electoral votes are needed to win the election, the probability of the Democratic ticket getting fewer than 270 electoral votes represents the probability of a Republican victory. And we can use the median Democratic electoral votes across one million hypothetical elections as a point estimate of the number for Democratic electoral votes.

Fig. 3 Comparing Forecasting Results from One Million Hypothetical Elections (Actual electoral votes: 306 Democratic, 232 Republican)



Data sources, November 3, 2020: https://www.predictit.org
https://projects.fivethirtyeight.com/2020-election-forecast/

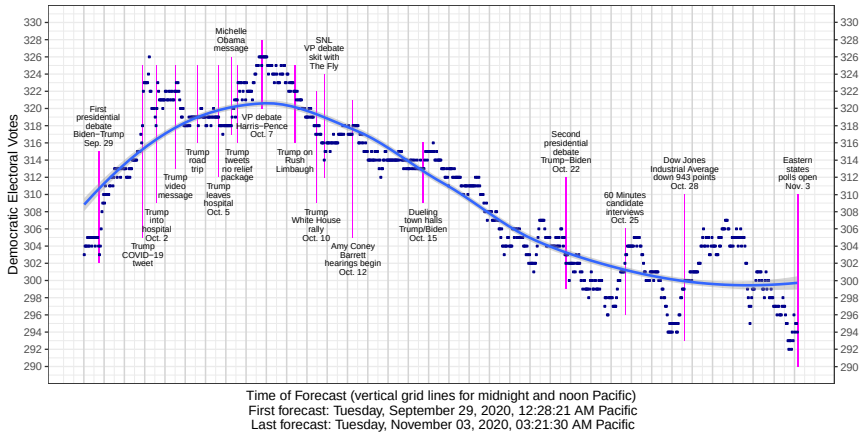
Figure 3 shows results from one million hypothetical elections using electoral market probabilities from four forecasting methods: a naïve 50-50 forecast, the 2016 presidential election baseline, the final FiveThirtyEight model prior, and the prediction-market-based forecast described in this paper (The Virtual Tout®). The last two of these forecasts were completed prior to the opening of the final day of voting, November 3, 2020. We can use medians from these simulations to evaluate forecast accuracy, given the actual electoral vote totals in 2020: 306 for the Democratic ticket of Biden-Harris versus 232 for the Republican ticket of Trump-Pence.

How far do forecasts for Democratic electoral votes differ from the actual Democratic electoral votes? The naïve 50-50 forecast gives equal numbers of electoral votes (269) to both Democratic and Republican tickets, yielding a forecast that is 37 votes below the actual number of Democratic electoral votes in 2020: $269 - 306 = -37$. Using forecasts based on the 2016 presidential election, we would obtain a forecast 33 votes below the actual number of electoral votes in 2020: $273 - 306 = -33$.

The FiveThirtyEight final model is less accurate than either of the baseline methods, yielding a forecast for Democratic electoral votes 42 votes higher than the actual number of electoral votes: $348 - 306 = 42$. Across the four methods being compared, the prediction-market-based model was clearly the most accurate. Prior to the final day of voting, Novermber 3, 2020, the forecast from The Virtual Tout® was 294 Democratic electoral votes, only 12 electoral votes lower than actual: $294 - 306 = -12$.

Prediction-market-based forecasts are not only more accurate than forecasts that depend on politcal polls. They are also useful in tracing the effects of campaign and news events. Figure 4 shows a time series of forecasts. Each plotted point represents the median Democratic electoral votes from one mil-

Fig. 4 Forecasts across Time Show Effects of Campaign and News Events



Data source: https://www.predictit.org, November 3, 2020

lion hypothetical elections. The figure shows forecasted Democratic electoral votes increasing dramatically in the days following the first Trump-Biden debate and Trump's tweet that he tested positive for COVID-19. There is a decline in Democratic prospects aligned with Trump's return to the campaign trail. In the last week of the contest, following a substantial decline in stock market averages, there is an increase in forecasted electoral votes for the Democratic ticket. This is followed by continued decline in forecasted Democratic electoral votes up to the final day of voting.

5 Conclusions

Prediction markets provide a meaningful alternative to forecasting methods based on historical observations and opinion polls. Past studies show that prediction markets provide accurate forecasts of election outcomes. Prediction markets are fast to respond to current events, making them a useful tool for tracing the progress of political campaigns.

Costs of prediction market data collection are minimal. With data from prediction markets in hand, we can easily simulate millions of hypothetical elections. And, most importantly, this study shows that prediction-market-based forecasts are more accurate than forecasts that rely on opinion polls.

6 About the Author

Thomas W. Miller is faculty director of the data science program at Northwestern University, author of six books about data science, and owner of Research Publishers LLC.

References

Abramowitz, Alan I. 2004. "Information Markets, Administrative Decision-making and Predictive Cost-Benefit Analysis". *The University of Chicago Law Review* 71 (3): 933–1020.

Abramowitz, Alan I. 2008. "Forecasting the 2008 Presidential Election with the Time-for-Change Model". *Political Science and Politics* 41 (4): 691–695.

Arnesen, Sveinung, and Ole Bergfjord. 2014. "Prediction Markets vs Polls—An Examination of Accuracy for the 2008 and 2012 Elections". *The Journal of Prediction Markets* 8 (3): 24–33.

Arrow, Kenneth J., Robert Forsythe, Michael Gorham, Robert Hahn, Robin Hanson, John O. Ledyard, Saul Levmore, Robert Litan, Paul Milgrom, Forrest D. Nelson, George R. Neumann, Marco Ottaviani, Thomas C. Schelling, Robert J. Shiller, Vernon L. Smith, Erik Snowberg, Cass R. Sunstein, Paul C. Tetlock, Philip E. Tetlock, Hal R. Varian, Justin Wolfers, and Eric Zitzewitz. 2008. "The Promise of Prediction Markets". *Science* 320 (5878): 277–295.

Berg, Joyce E., and Thomas A. Rietz. 2006. "The Iowa Electronic Markets: Stylized Facts and Open Issues". In *Information Markets: A New Way of Making Decisions*, eds. Robert W. Hahn and Paul C. Tetlock, 142–169. Washington, DC: AEI-Brookings Joint Center for Regulatory Studies.

Berg, Joyce, Robert Forsythe, Forrest Nelson, and Thomas Rietz. 2008. "Results from a Dozen Years of Election Futures Markets Research". In *Handbook of Experimental Economics Results*, eds. Charles R. Plott and Vernon L. Smith, Vol. 1, 742–751. Amsterdam: North Holland/Elsevier. Working paper from March 2003 available at https://iemweb.biz.uiowa.edu/research/papers/BergForsytheNelsonRietz_2008.pdf.

Blackwell, Calvin, and Robert Pickford. 2011. "The wisdom of the few or the wisdom of the many? An indirect test of the marginal trader hypothesis". *Journal of Economics and Finance* 35: 164–180.

Brier, Glenn W. 1950. "Verification of Forecasts Expressed in Terms of Probability". *Monthly Weather Review* 78 (1): 1–3.

Campbell, James E. 2014. "Issues in Presidential Election Forecasting: Election Margins, Incumbency, and Model Credibility". *Political Science and Politics* 47 (2): 301–303.

Casati, B., L. J. Wilson, D. B. Stephenson, P. Nurmi, A. Ghelli, M. Pocernich, U. Damrath, E. E. Ebert, B. G. Brown, and S. Mason. 2008. "Forecast Verification: Current Status and Future Directions". *Meterorological Applications* 15: 3–18.

Crane, Harry. 2018. "Polls, Punditry, and Prediction Markets: An Assessment of Election Forecasting". *Researchers.One*. Retrieved from the World Wide Web on October 31, 2020, at https://researchers.one/articles/polls-pundits-or-prediction-markets-an-assessment-of-election-forecasting/5f52699b36a3e45f17ae7d8c/v1.

Crane, Harry, and Darrion Vinson. 2020. "Models vs. Markets: Forecasting

the 2020 U.S. Election". *Researchers.One*. Retrieved from the World Wide Web on October 31, 2020, at https://researchers.one/articles/Models-vs-Markets-Forecasting-the-2020-US-election/97212967bc124e2bf3af31a5/v1.

Crossley, Archibald M. 1937. "Straw Polls in 1936". *Public Opinion Quarterly* 1 (1): 24–35.

Economist, Andrew Gelman, and Merlin Heidemanns. 2020. "How The Economist Presidential Forecast Works". *The Economist*. Retrieved from the World Wide Web on September 9, 2020 at https://projects.economist.com/us-2020-forecast/president/how-this-works Daily forecast updates provided at https://projects.economist.com/us-2020-forecast/president.

Erikson, Robert S., and Christopher Wlezien. 2008a. "Are Political Markets Really Superior to Polls as Election Predictors?". *The Public Opinion Quarterly* 72 (2): 190–215.

Erikson, Robert S., and Christopher Wlezien. 2008b. "Leading Economic Indicators, the Polls, and the Presidential Vote". *Political Science and Politics* 41 (4): 703–707.

Erikson, Robert S., and Christopher Wlezien. 2012. "Markets vs. Polls as Election Predictors: An Historical Assessment". *Electoral Studies* 31: 532–539.

Fair, Ray C. 1978. "The Effect of Economic Events on Votes for President". *Review of Economics and Statistics* 60: 159–173.

Fair, Ray C. 1996. The Effect of Economic Events on Votes for President: 1992 Update. Retrieved from the World Wide Web on September 10, 2020, at https://fairmodel.econ.yale.edu/RAYFAIR/pdf/1996A300.PDF.

Fair, Ray C. 2011. *Predicting Presidential Elections and Other Things*, 2nd edn. Palo Alto, CA: Stanford Economics and Finance.

Fama, Eugene F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work". *Journal of Finance* 25: 383–417.

Fama, Eugene F. 1991. "Efficient Capital Markets: II". *Journal of Finance* 46: 1575–1617.

FiveThirtyEight. 2020a. "FiveThirtyEight's Pollster Ratings" FiveThirtyEight. Retrieved from the World Wide Web on September 7, 2020, at https://projects.fivethirtyeight.com/pollster-ratings/, with pollster data available on GitHub at https://github.com/fivethirtyeight/data/tree/master/pollster-ratings.

FiveThirtyEight. 2020b. "Forecasting Models" FiveThirtyEight. Retrieved from the World Wide Web on September 10, 2020, at https://fivethirtyeight.com/tag/forecasting-models/.

Forsythe, Robert, Forrest Nelson, George R. Neumann, and Jack Wright. 1992. "Anatomy of an Experimental Political Stock Market". *American Economic Review* 82: 1142–1161.

Gelman, Andrew, and Gary King. 1993. "Why Are American Presidential Election Campaign Polls So Variable When Votes Are So Predictable". *British Journal of Political Science* 23 (4): 409–451.

Gelman, Andrew, Jessica Hullman, Christopher Wlezien, and George Elliott Morris. 2020. "Information, Incentives, and Goals in Election Forecasts". *Judgment and Decision Making* 15 (5): 863–880.

Gneiting, Tilmann, and Adrian E. Raftery. 2007. "Strictly Proper Scoring Rules, Prediction, and Estimation". *Journal of the American Statistical Association* 102 (477): 359–378.

Groves, Robert M., Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, and Eleanor Singer. 2009. *Survey Methodology*, 2nd edn. New York: Wiley.

Hahn, Robert W., and Paul C. Tetlock, eds. 2006. *Information Markets: A New Way of Making Decisions*. Washington, DC: AEI-Brookings Joint Center for Regulatory Studies.

Hayek, Friedrich A. 1945. "The Use of Knowledge in Society". *American Economic Review* 35 (4): 519–530.

Hillygus, D. Sunshine. 2011. "The Evolution of Election Polling in the United States". *The Public Opinion Quarterly* 75 (5): 962–981.

Holbrook, Thomas M. 2008. "Economic Considerations and the 2008 Presidential Election". *Political Science and Politics* 42 (3): 473–478.

Holbrook, Thomas M. 2012. "Incumbency, National Conditions, and the 2012 Presidential Election". *Political Science and Politics* 45 (4): 640–643.

Honomichl, Jack J. 1990. *Honomichl on Marketing Research*. Lincolnwood, IL: National Textbook Company.

Horn, Christian Franz, Bjoern Sven Ivens, Michael Ohneberg, and Alexander Brem. 2014. "Prediction Markets: A Literature Review 2014 Following Tziralis and Tatsiopoulos". *The Journal of Prediction Markets* 8 (2): 89–126.

ICPSR. 2020. American National Election Studies (ANES) Interuniversity Consortium for Political and Social Research. Retrieved from the World Wide Web on September 8, 2020, at
https://www.icpsr.umich.edu/web/pages/,
with ANES data available at
https://www.icpsr.umich.edu/web/ICPSR/search/studies?q=anes.

Jacobs, Lawrence R., and Melanie Burns. 2004. "The Second Face of the Public Presidency: Presidential Polling and the Shift from Policy to Personality Polling". *Presidential Studies Quarterly* 34 (3): 536–556.

Jacobs, Lawrence R., and Robert Y. Shapiro. 1995. "The Rise of Presidential Polling: The Nixon White House in Historical Perspective". *Public Opinion Quarterly* 59 (2): 163–195.

Jacobs, Lawrence R., and Robert Y. Shapiro. 2005. "Polling, Politics, Media, and Election Campaigns". *Public Opinion Quarterly* 69 (5): 635–641.

Jerôme, Bruno, and Véronique Jerôme-Speziari. 2012. "Forecasting the 2012 US Presidential Election: Lessons from a State-by-State Political Economy Model". *Political Science and Politics* 45 (4): 663–668.

Kou, S. G., and Michael E. Sobel. 2004. "Forecasting the Vote: A Theoretical Comparison of Election Markets and Public Opinion Polls". *Political Analysis* 12 (3): 277–295.

Leigh, Andrew, and Justin Wolfers. 2006. "Competing Approaches to Forecasting Elections: Economic Models, Opinion Polling and Prediction Markets".

The Economic Record 82 (258): 325–340.

Lewis-Beck, Michael S., and Charles Tien. 2008. "Forecasting presidential elections: When to change the model". *International Journal of Forecasting* 24 (2): 227–236.

Lichtman, Allan J. 2008. "The Keys to the White House: An Index Forecast for 2008". *International Journal of Forecasting* 24: 301–309.

Lichtman, Allan J. 2020. *Predicting the Next President: The Keys to the White House*, reprint edn. Lanham, MD: Rowman & Littlefield.

Luckner, Stefan, Jan Schröder, Christian Slamka, Bernd Skiera, Martin Spann, Christof Weinhardt, Andreas Geyer-Schulz, and Markus Franke. 2012. *Prediction Markets: Fundamentals, Designs, and Applications*. Germany: Gabler Verlag/Springer.

Manski, Charles F. 2006. "Interpreting the Predictions of Predictions Markets". *Economics Letters* 91 (3): 425–429.

Miller, Thomas W. 2001. "Can We Trust the Data of Online Research?". *Marketing Research* Summer: 26–32. Reprinted as "Online Results are a Mixed Bag," *Marketing News*, September 24, 2001, 20–25.

Miller, Thomas W. 2008. *Research and Information Services: An Integrated Approach for Business*. Manhattan Beach, CA: Research Publishers LLC.

Miller, Thomas W., and Peter R. Dickson. 2001. "On-Line Market Research". *International Journal of Electronic Commerce* 5 (3): 139–167.

Murphy, Allan H. 1993. "What is a Good Forecast? An Essay on the Nature of Goodness of Weather Forecasting". *Weather and Forecasting* 8: 281–293.

Oliven, Kenneth, and Thomas A. Rietz. 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.

PredictIt. org. 2020. Which party will win the Electoral College? PredictIt.org. Retrieved from the World Wide Web on August 30, 2020, at https://www.predictit.org/.

Rhode, Paul W., and Koleman S. Strumpf. 2004. "Historical Presidential Betting Markets". *Journal of Economic Perspectives* 18 (2): 127–141.

Rosenstone, Steven J. 1983. *Forecasting Presidential Elections*. New Haven, CT: Yale University Press.

Rothschild, David. 2009. "Forecasting Elections: Comparing Prediction Markets, Polls, and Their Biases". *Public Opinion Quarterly* 73 (5): 895–916.

Selten, Reinhard. 1998. "Axiomatic Characterization of the Quadratic Scoring Rule". *Experimental Economics* 1: 43–62.

Silver, Nate. 2012. *The Signal and the Noise: Why So Many Predictions Fail—But Some Don't*. New York: Penguin.

Stix, Gary. 2008. "When Markets Beat the Polls". *Scientific American* 298 (3): 38–45.

Sunstein, Cass R. 2005. "Group Judgments: Statistical Means, Deliberation, and Prediction Markets". *New York University Law Review* 80: 962–1049.

Sunstein, Cass R. 2006a. "Deliberation and Information Markets". In *Information Markets: A New Way of Making Decisions*, eds. Robert W. Hahn and Paul C. Tetlock, 67–100. Washington, DC: AEI-Brookings Joint Center

for Regulatory Studies.

Sunstein, Cass R. 2006b. *Infotopia: How Many Minds Produce Knowledge*. New York: Oxford University Press.

Surowiecki, James. 2005. *The Wisdom of Crowds*. New York: Anchor.

Tetlock, Philip E. 2017. *Expert Political Judgment: How Good Is It? How Can We Know?*, new edn. Princeton, N.J.: Princeton University Press.

Theiss-Morse, Elizabeth A., Michael W. Wagner, William H. Flanigan, and Nancy H. Zingale. 2018. *Political Behavior of the American Electorate*, fourteenth edn. New York: Anchor.

Traugott, Michael W. 2014. "Public Opinion Polls and Election Forecasting". *Political Science and Politics* 47 (2): 342–344.

Tziralis, Georgios, and Ilias Tatsiopoulos. 2007. "Prediction Markets: An Extended Literature Review". *The Journal of Prediction Markets* 1 (2): 75–91.

Weisberg, Herbert F. 2005. *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago: University of Chicago Press.

White, Daniel. 2015. "Here's Why Gallup Won't Poll the 2016 Election". *Time*. Retrieved from the World Wide Web on September 21, 2020 at https://time.com/4067019/gallup-horse-race-polling/.

Williams, Leighton Vaughan, ed. 2005. *Information Efficiency in Financial and Betting Markets*. Cambridge, UK: Cambridge University Press.

Wolfers, Justin, and Eric Zitzewitz. 2004. "Prediction Markets". *Journal of Economic Perspectives* 18 (2): 107–126. https://pubs.aeaweb.org/doi/pdfplus/10.1257/0895330041371321.

Wolfers, Justin, and Eric Zitzewitz. 2006a. "Five Open Questions about Prediction Markets". In *Information Markets: A New Way of Making Decisions*, eds. Robert W. Hahn and Paul C. Tetlock, 13–36. Washington, DC: AEI-Brookings Joint Center for Regulatory Studies.

Wolfers, Justin, and Eric Zitzewitz. 2006b. "Interpreting Prediction Market Prices as Probabilities" Federal Reserve Bank of San Francisco. Retrieved from the World Wide Web on September 7, 2020, at https://www.frbsf.org/economic-research/files/wp06-11bk.pdf.

Appendix 1. Presidential Election Data from Prediction Markets

state	votes	dem	rep	demprob	repprob
AK	3	13	87	0.130	0.870
AL	9	4	95	0.040	0.960
AR	6	6	94	0.060	0.940
AZ	11	55	47	0.539	0.461
CA	55	95	6	0.941	0.059
C0	9	87	13	0.870	0.130
CT	7	95	6	0.941	0.059
DC	3	97	3	0.970	0.030
DE	3	96	5	0.950	0.050
FL	29	53	48	0.525	0.475
GA	16	33	67	0.330	0.670
HI	4	97	4	0.960	0.040
IA	6	28	72	0.280	0.720
ID	4	5	95	0.050	0.950
IL	20	95	7	0.931	0.069
IN	11	8	92	0.080	0.920
KS	6	7	92	0.071	0.929
KY	8	5	94	0.051	0.949
LA	8	6	94	0.060	0.940
MA	11	96	5	0.950	0.050
MD	10	96	5	0.950	0.050
ME	2	83	19	0.814	0.186
ME-01	1	95	7	0.931	0.069
ME-02	1	29	72	0.287	0.713
MI	16	71	32	0.689	0.311
MN	10	68	32	0.680	0.320
MO	10	11	90	0.109	0.891
MS	6	6	94	0.060	0.940
MT	3	10	91	0.099	0.901
NC	15	44	57	0.436	0.564
ND	3	4	96	0.040	0.960
NE	2	6	93	0.061	0.939
NE-01	1	6	95	0.059	0.941
NE-02	1	55	46	0.545	0.455
NE-03	1	2	98	0.020	0.980
NH	4	69	32	0.683	0.317
NJ	14	94	8	0.922	0.078
NM	5	88	11	0.889	0.111
NV	6	80	21	0.792	0.208
NY	29	94	7	0.931	0.069
OH	18	37	66	0.359	0.641
OK	7	4	96	0.040	0.960
OR	7	94	9	0.913	0.087
PA	20	64	39	0.621	0.379
RI	4	96	4	0.960	0.040
SC	9	9	91	0.090	0.910
SD	3	5	97	0.049	0.951
TN	11	5	95	0.050	0.950
TX	38	25	76	0.248	0.752
UT	6	8	93	0.079	0.921
VA	13	87	13	0.870	0.130
VT	3	98	5	0.951	0.049
WA	12	94	7	0.931	0.069
WI	10	58	45	0.563	0.437
WV	5	4	96	0.040	0.960
WY	3	4	96	0.040	0.960

Data source: https://www.predictit.org, August 28, 2020