Combining Forecasts for Elections: Accurate, Relevant, and Timely

David Rothschild[a]

[a] *Microsoft Research, New York, NY, USA*

Abstract

This paper increases the efficiency and understanding of forecasts for Electoral College and senatorial elections by generating, and then combining, forecasts based on voter intention polling, fundamental data, and prediction markets. The paper addresses the most relevant outcome variable, probability of victory in state-by-state elections, while also solving for the traditional outcomes, and ensuring that the forecasts easily update continuously during the course of the main election cycle. Attempting to maximize those attributes and accuracy, I create efficient forecasts, for each of these three types of raw data, with innovations in aggregating the data, and then correlating the aggregated data with outcomes. The paper demonstrates that all three data types provide significant and meaningful contributions towards election forecasts. Varied stakeholders, including researchers, election investors, and election workers, can all benefit from the efficient combined forecasts defined in this paper; the forecast is tested and excels out-of-sample during the 2012 elections.

Keywords:
election forecasting, surveys, econometric models, prediction markets, combining forecasts, probability forecasting

1. Introduction

Polling data has been the most prominent component of election forecasts for decades. From 1936 to about 2000, it was standard in both the academic and popular press to utilize just the raw data, the results of individual voter intention polls, as an implicit forecast of an election. By 2004 poll aggregation became common on the internet. Although aggregated polls provide both stability and accuracy relative to individual poll results, aggregated polls are meant to be a closer approximation, relative to individual poll results, of what an election would look like if it was suddenly held on that day, not an expectation of what will happen on Election Day. By 2008 some websites, run by a mix of academics and non-academics, finally began publishing poll-based forecasts (i.e., forecasts derived from aggregating raw polls and subsequently translating that into a forecast of the election outcome). Further,

Email address: DavidMR@Microsoft.com (David Rothschild)

I thank the editors and reviewers for their extremely helpful criticism, and Sunshine Hillygus and Jolie Martin for their insightful comments.

they shifted the outcome variable to the probability of victory in the Electoral College or senatorial elections versus the standard expected vote shares of the national popular vote.

The need to transform raw polling data into a forecast is conclusive in the literature. Campbell (2008) clearly illustrates the anti-incumbency bias, whereby incumbents have lower polling values than election results, and the fading of early leads in polls, whereby election results are tighter than polling numbers. Erikson and Wlezien (2008a) shows that translating raw polling data into a forecast makes it more accurate for both estimated vote share and probability of victory. Rothschild (2009) improves on Erikson and Wlezien (2008a) by aggregating the daily polls over time, eliminating noisy daily fluctuations, and then translating them into a forecast. Yet, Rothschild (2009) designed its poll-based forecast to be the most accurate forecast using the same general model as Erikson and Wlezien (2008a), leaving open the possibility of creating even more accurate transformations through more advanced models of aggregation and subsequent translation of the polling data into forecasts. The most recent advances in creating forecasts from polls have been in the methods of aggregation, including eliminating poll company specific effects and combining the snapshot for any given state with other state polls and national polls. In this paper, when available, we use both the most transparent and efficient method possible, without these further steps, and Stanford's Simon Jackman's interpretation of these steps, as made available through Pollster.com (Jackman, 2005).

There is a massive literature on modeling fundamental data, most models are not useful as forecasts, but rather explain the correlation between different variables and election outcomes. These models use a variety of economic and political indicators such as: past election results, incumbency, presidential approval ratings, economic indicators, ideological indicators, biographical information, policy indices, military situations, and facial features of the candidates. There is a substantial list of models in Hummel and Rothschild (2013), however, there are several reasons they are generally not useful in creating forecasts. First, many models are difficult to duplicate, such as Armstrong et al. (2010) which utilizes pictures of the candidates. Second, many models incorporate pre-election polls or other late arriving data; for example Lock and Gelman (2010) uses a model that cannot be resolved until October of the election year. These types of models are designed more towards understanding the correlation of fundamental data and elections outcomes, than forecasting the election during the cycle. Third, most fundamental data models forecast just the presidential national popular vote; examples include Abramowitz (2004, 2008). This is a serious issue not just because it is not the ideal outcome variable, but that means there is extremely limited identification in just one outcome every four years. Fourth, Klarner (2008) pushed the literature forward into the realm of earlier state-by-state forecasts, but still incorporated early polling into the model. In order to compare the value of the different data sources, it is crucial to use models that are purely one data source. Without any polling data, improving on the variable choice and range of data, the model presented by Hummel and Rothschild (2013) has much smaller errors than Klarner (2008) and could be put to use by June 15 of the election year. Thus, I utilized Hummel and Rothschild (2013) exclusively as the fundamental model for this paper because it is the most accurate, state-by-state fundamental model for Electoral College and senatorial election, can be executed early in the cycle, and it excludes voter intention polling data. The out-of-sample errors for Hummel and Rothschild (2013) are smaller than the within-sample errors for the most widely circulated state-by-state fundamental models, including Klarner's most recently updated model (Klarner, 2012).[2]

The modern history of prediction markets is not as long as the other two data sources. The Iowa Electronic Market launched the modern era of prediction markets in 1988, introducing a winner-takes-all market in 1992. This type of market trades binary options which pay, for example, \$10 if the chosen candidate wins and \$0 otherwise. Thus, an investor who pays \$6 for a Democrat to Win stock, and holds the stock through Election Day, earns \$4 if the Democrat wins and loses \$6 if the Democrat loses. In that scenario, if there are no transaction or opportunity costs, the investor should be willing to pay up to the price that equals her estimated probability of the Democrat winning the election. The market price is the value at which, if a marginal investor were willing to buy above it, investors would sell the contract and drive the price back down to that market price (and vice-versa if an investor were willing to sell below it); thus, the price is an aggregation of the subjective probability beliefs of all investors.

Scholars have found that prediction market prices can create more accurate forecasts than poll-based forecasts in the last few cycles (Berg et al., 2008; Rothschild, 2009) and in historical elections (Rhode and Strumpf, 2004); but, like polling and fundamental data, prediction market prices benefit from a transformation from raw data into a forecast, especially due to the favorite-longshot bias. Berg et al. (2008) concludes that raw prediction market prices are more accurate forecasts of vote share than raw polling data. Erikson and Wlezien (2008a) challenges this finding by comparing raw prediction market prices with properly translated poll-based forecasts; this confirmed by Rothschild (2009). Yet, Wolfers and Zitzewitz (2006) highlights the transaction and opportunity costs of investing in prediction markets, Manski (2006) describes how investors in prediction markets behave as if they were risk-loving, and Snowberg and Wolfers (2010) concludes that there are systematic mis-perceptions of probability stemming from prospect theory; the result of those three papers is the favorite-longshot bias for prediction market prices. One hundred days before the election, if an investor believes the Republican candidate has a 95 percent chance of winning, there are three reasons for her to bid less than \$0.95 for a contract that pays out \$1.00 if the candidate wins. First, with limited liquidity in the market (i.e., not enough traders and money in the market for all traders to always make their most efficient purchases) she may have to hold the contract until Election Day, incurring an opportunity cost. Second, she will incur some transaction costs when she buys and sells the contract or when it expires. If the opportunity cost is \$0.02 and the transaction cost is \$0.03, then she would not bid more than \$0.90 in

[2]Klarner (2012) drops voter intention polling, which Klarner used in early versions of the model, but that was released after the initial running of the model for and circulation of this paper, as I wanted to ensure that 2012 would be completely out-of-sample; that change brought Klarner closer to Hummel and Rothschild (2013), which was already a widely circulated working paper at that time. Still, while Hummel and Rothschild (2013) had similar errors for estimated vote share in 2012 to Klarner (2012) for the Electoral College, it had significantly smaller errors, nearly a full point on average, for the senatorial elections.

order to break even in expectation. Third, investors behaving as if they were risk loving gain greater expected utility from buying a longshot than a favorite, all else being equal. Thus, even though the unbiased, risk-neutral market price would be \$0.95, investors do not bid all the way to the biased, risk neutral amount of \$0.90 for a favored candidate, but may move the market to only \$0.85 or less. Rothschild (2009) corrected for this bias with a transformation suggested by Leigh et al. (2007) and proved that corrected prediction market data generates more accurate forecasts than aggregated and corrected polling data.[3] Still, Rothschild (2009) makes no attempt to examine or improve upon the transformation suggested by Leigh et al. (2007), which was calibrated on a different type of prediction market data; this paper takes the logical next steps.[4]

There is a rich history of combining data for forecasts, in many domains, including elections, to gain accuracy; nothing in the literature does state-by-state with all three types of data. Clemen (1989) provides an early overview across domains. Graefe et al. (2012) concludes there are benefits in accuracy to combining many data types for forecasts of the vote share of the national popular vote for president. Yet, that paper suggests the simple procedure of using even weights at all points in the cycle, uses simplified forecasts from the various data types, and only examines the national popular vote. Erikson and Wlezien (2008b) combines economic indicators and intention polls to forecast the national popular vote, but that paper does not investigate prediction markets, state-by-state elections, or probability of victory. Erikson and Wlezien (2012) attempts to empirically demonstrate that raw prediction market prices provide no information beyond polling information, but that paper again just looks at a few national popular votes; that is the wrong outcome variable for prediction markets. There is no expectation that vote share is well identified by prediction markets that generally trade on the probability of the election winner (i.e., if a candidate is poised to win big, the candidate should trade near \$1.00 per \$1.00 of payout for the candidate's victory regardless of whether it is a 5 or 10 point expected victory).

A few recent Bayesian papers, including Montgomery et al. (2012), Linzer (2013), and Lock and Gelman (2010), combine fundamental data and voter intention polling, but there are some key differences in how I approach the data. First, Linzer (2013) for example, starts with a fundamental-based forecast for each state and lets the forecasts update as the new polls arrive. I translate the separate data types into forecasts independently, rather than together, so that I can compare their informational value. This is not a criticism of Linzer (2013) and similar papers, but essential to the goal of this paper in comparing the different data types separately. Second, I allow the relative value of different data types in my final forecast to fluctuate by day while, for example, Linzer (2013) focuses on the certainty of the data by race. This paper provides a simpler model and highlights some key points about shifting information over time, but again, this is allows me to optimize

[3]Interestingly, Erikson and Wlezien (2008a) recognized the existence of the favorite-longshot bias in prediction markets, but then did not correct for it.

[4]Leigh et al. (2007) was calibrated on national presidential elections, while Rothschild (2009) utilizes state-by-state Electoral College and senatorial elections from recent cycles.

over different outcomes than Linzer (2013) and similar papers, and it not a criticism of that work. Third, these papers exclude prediction market data. Prediction market theory suggests that prediction market prices should include the information from voter intention polls and fundamental models already, but the literature lacks any empirical work on the subject. This paper uniquely provides a clean comparison of the three data types in the key setting of state-by-state elections.

This paper defines an efficient election forecast with three attributes geared towards maximizing returns for both election workers and researchers: state-by-state probability of victory, regularly updating, and accurate. Historically election forecasts estimated vote share rather than probability of victory for two key reasons: academic literature focuses on incremental improvements on historical forecasts and estimated vote share is the historical standard, and observers frequently interpret raw polls as a naïve estimations of vote share, making it the simplest rubric. Expected vote share is extremely important for election workers, especially broken down by targetable demographics. But, most stakeholders care about shifts in the expected vote share insofar as they affect the probability of victory; when judging the impact of a debate on the outcome of the election or deciding on which race to invest, probability of victory is a more efficient metric. Historically election forecasts estimated national outcomes, rather than state-by-state, but the U.S. elects its president through the Electoral College, not the popular vote.[5] Further, any forecast that exclusively utilizes national outcomes has a serious issue with identification, as national elections occur so infrequently; state-by-state elections can correlate, but they provide extra identification. There is a revealed preference for probability of victory on the state-by-state level as the main popular forecasting sites from FiveThirtyEight to Princeton Election Consortium report probability of victory in the Electoral College as their main forecast (and barely mention expected vote share, especially the national popular vote).[6] There are increasing instances of probability of victory utilized in academia, such as Lewis-Beck and Rice (1992) and state-by-state outcomes, such as Linzer (2013), yet the influential PS: Political Science's special edition journal in 2012, "Forecasting the 2012 American National Election", had probability of victory for the national popular vote along with expected national popular vote Campbell (2012) and the Iowa Electronic Market still uses victory in the national popular vote for its main market, not Electoral College.

Historically forecasts were likely to update with other newly released information, such as a new poll, or right before the event when all resources have been allocated and major events have already passed; but, forecasts are more valuable further in advance of the event and if they update regularly, so they can exist and be fresh when stakeholders can still reallocate investments to more efficient uses.[7] For researchers, the earlier the forecast, the more events they capture and time-granular forecasts are necessary to study the effect of small events

[5] On four separate occasions the winner of the national popular vote lost the Electoral College!

[6] FiveThirtyEight accounted for 20% of the New York Times traffic in the lead-up to the 2012 election http://money.cnn.com/2012/11/07/news/companies/nate-silver-election/index.html.

[7] According to the Center for Responsive Politics, \$6.3 billion was spent on the 2012 election with \$2.6 billion going to the presidential election alone: http://www.opensecrets.org/bigpicture.

that occur during the lead up to an election; they allow researchers to isolate the effect of debates on outcomes or study the effects of elections on other events, as Snowberg et al. (2006) did for elections and the economy. Accuracy includes the smallest errors, the most efficient calibration, and out-of-sample robustness (i.e., explains the future, not describes the past). These extra checks, beyond the simple error, help off-set the difficulty in judging forecasts, where there is an uncertain ground truth on all days except the last (i.e., we do not know how the election would have turned out on any day other than the final day).

This paper demonstrates that all three data types, polls, fundamentals, and prediction markets, should be part of an efficient election forecast. The academic literature is clear that combining data is generally very effective in increasing accuracy (Clemen, 1989). Yet, overall, three related, but largely non-intersecting, literatures persist for each data type. The closest work to this paper is Linzer (2013), which models state-by-state forecasts for the Electoral College and utilizes polling and fundamental data, but among other differences, there is no literature that tests prediction market prices as well for these efficient outcomes. The forecast created in this paper appreciates that information provided by the three different forecasts types shifts during the studied timeframe; 130 days out, the most efficient combined forecast averages the forecasts from all three data types, but the fundamental data's unique information decreases, so that the Election Day forecast averages just polling and prediction market data. Varied stakeholders, including researchers, election investors, and election workers, can all benefit from the efficient combined forecast for Electoral College or senatorial elections defined in this paper; the forecast is tested and excels out-of-sample during the 2012 elections.

2. Data

This paper calibrates the combined forecast with data from 202 races that span four election cycles. Electoral College races are included for 2004 and 2008, and senatorial elections are included for 2006, 2008, and 2010. In this sample there are 100 Electoral College races, or 50 for each cycle.[8] There are 33 senatorial elections in 2006, 35 in 2008, and 37 in 2010. I exclude three senatorial elections from the calibration because they are outliers, but include them in any derived forecasts.[9] The 33 senatorial elections and 50 Electoral College races of 2012 serve as an out-of-sample test for the combined forecast.[10] Data from the 2000 Electoral College races and 2004 senatorial elections help calibrate the poll-based forecast.

The polling data is as complete and accurate as possible. I gathered the voter intention poll from all state-by-state polls listed on either PollingReport.com, Pollster.com, and RealClearPolitics.com; I used several sites to ensure that I have as many public polls as possible.

[8] I exclude Washington DC, which votes Democratic in all presidential elections with 100 percent likelihood.

[9] There were three unique three-way senate elections in the sample: Connecticut in 2006 and Alaska and Florida in 2010.

[10] This paper was circulated publicly prior to the 2012 election to ensure that the forecasts can be considered ex-ante.

Pollster.com itself creates a rolling average that incorporates research such as (Jackman, 2005) and (Pickup and Johnston, 2005) on how to eliminate possible house biases in polls and draw updates from other state and national polls. Neither I, nor Pollster.com, have these aggregated averages historically. To create the historical polling snapshots I utilized the most advanced, transparent aggregation methods, but I do not go back and correct for house bias historically, because any attempt would suffer from a potential look-ahead bias and the impact would be minimal. I ensure this is not an issue by testing my 2012 out-of-sample results with both my method of aggregating polls and Pollster.com aggregated average (when available); all results are robust to either.

The fundamental model follows the procedure outlined in Hummel and Rothschild (2013). The authors created the models, separately for Electoral College and senatorial elections, using data from the following categories: election results, incumbency, presidential approval ratings, economic indicators, ideological indicators, and biographical information about the candidates. The model is calibrated on Electoral College data from 1972-2008 and senatorial data from 1976-2010. All of the data for the fundamental model is publicly available on government websites, with exception of presidential approval ratings which are gathered at Gallup and Pollster.com (for later years).

The prediction market data is all gathered from Intrade in real-time. It sells contracts for all candidates that are worth either \$10 if that candidate wins or \$0 if that candidate loses. I have marginal order book for all elections from 2004, 2006, 2008, 2010, and 2012 saved in 10 minute intervals throughout the election cycle. For each candidate the data includes: the price people are willing to buy (bid), sell, and the last price sold, along with the volume of trade. Intrade is the most liquid market for Electoral College and senatorial elections. Where the data is available, results are similar Betfair, the world's largest prediction market as well, but I only have Betfair data from 2010 and 2012 elections.

In order to approximate a random draw of demand for information, there is one observation per forecast type for each day. For the polling data, I date polls by the last day they are in the field; if anything, this is biased in favor of the relevancy of polls, as polls are frequently released at least one day after they leave the field. For the fundamental model used in this paper, that value is the same for the entire cycle. The model shows that there is little to no added information in including late breaking fundamental data. For prediction market data, I use the average of the bid and ask at noon on the noted day.

I analyze all data for the time period between 0 and 130 days before the election; 130 days generally falls in late June of the election year and it is when the fundamental model is realized. The standard start of the campaign season in the United States is Labor Day, which falls roughly 60-65 days before the election, so this allows me to review a sizable amount of time in both the before and after period.

All three forecasts types do not occur on all days. First, there is a liquidity issue for prediction markets early in the cycle, where the lack of meaningful trading can lead to some imprecise pricing, although the markets all exist. Second, some senatorial elections never have major polls in the entire election cycle, while many do not start having polls until midway through the timeframe. Third, I have a slight data issue where for 10 random days in 2008 my program did not record prediction market data.

There are three additional dimensions of the data to consider beyond data type: election type, days before the election, and the certainty of the raw data. First, where the identification exists, I create different forecast models for the Electoral College and senatorial elections. Second, some parameters vary depending on the days until the election. Finally, some parameters vary depending on the certainty of the raw data. This can be important, because, for example, a prediction market price near 50 percent may correlate with expected outcomes differently than a price near 95 percent, or a poll that gives a 2 point lead may behave differently than a poll with a 10 point lead.

The first step in this process is to create the most efficient possible forecasts for estimated vote share and probability of victory for all three forecast types. I could go through the process of aggregating all information simultaneously into one combined forecasts, but there are two major advantages to first creating three separate forecasts and then combining them. First, this puts all of the data on the same scale, allowing us to understand how the information mixes in the combined forecast. Second, not all data is available at all times, so it is a plus for forecasting to have both the separate and combined versions.

This paper emphasizes probability of victory over estimated vote share, but estimated vote share is included to provide historical comparisons.

3. Estimation Strategy/ Results in Creating Separate Forecasts

The fundamental model Hummel and Rothschild (2013) utilizes an OLS regression to predict the expected vote share and a probit regression to predict the probability of victory. The models are capable of creating forecasts prior to 130 days before the election for both types of elections. I encourage readers to review Hummel and Rothschild (2013) if they wish to recreate the fundamental model, but for reference the Electoral College model has the following coefficients for Democratic vote share: 0.41 on (presidential approval - 42)*incumbency, -1.88 on two or more terms*incumbency, 0.72 on state vote four years ago - national vote, 0.12 on state vote four years ago - national vote, 0.21 on change in state income from 9th to 13th quarter of term, -0.02 on sum of ACU rankings for senators - average sum of ACU senators, 0.08 on change in %Dems in lower house of state legislature, 4.83 on home state if less than 10 million in population, -3.04 on home state from last cycle if less than 10 million population, and 47.37 as a constant. The senatorial model has the following coefficients for Democratic vote share: 0.13 on (presidential approval - 50)*presidential incumbency, 12.82 on incumbency, -2.80 on midterm*presidential incumbency, 0.36 on last presidential vote - national vote, 0.11 on state vote six years ago - national vote, 0.21 on change in state income from 9th to 13th quarter of term*presidential incumbency, 0.09 on Reps ACU rating - 74, 0.08 on change in %Dems in lower house of state legislature, between 3.87 and 10.39 on previous job of senatorial candidates, and 46.55 as a constant.[11]

[11] I exclude all coefficients that do not apply to 2012 and forward. I have included the coefficients for the probit, to create probability of victory, in the Appendix.

The voter intention polling and prediction market models follow the procedures created in Rothschild (2009), but incorporate some major advances beyond that paper. I update the debiasing techniques (i.e., the translation of the aggregated raw data into a forecast) used for both forecast types. The polling debiasing technique varied by election type and days before the election in Rothschild (2009); this paper also allows variation by the size of the lead in the two party vote share. For prediction markets, that article had a one-size-fits-all debiasing, but I examine the value of allowing it to fluctuate along each of three dimensions: election type, days before the election, and the certainty of the raw pricing (i.e., how close the price is to \$0.50 rather than \$0.00 or \$1.00 for contracts that pay out \$1.00).

The first step in creating a voter intention poll-based forecast is to create a snapshot of the estimated two-party vote share of the two candidates (i.e., candidate one's support divided by the sum of candidate one and two's support) if the election were held that day. This aggregation helps smooth the random fluctuation that occurs in raw daily polls. The method creates a linear regression of all polls up to that day, and the snapshot is the trend of that regression.[12] The data is de-meaned around zero, so that the two-party vote share runs from -0.5 to 0.5, where the leading candidate is above zero and the losing candidate is below zero. All vote shares are in terms of the state's incumbent (i.e., the winner of the state's electoral votes four years before and the state's senatorial election six years before).

The second step in creating a poll-based forecast is to create an estimated vote share for Election Day. To create the estimated vote share I regress the final vote share on the poll for each day before the election in previous election years:

$$(1) V_r = \alpha + \beta_1 S_r + \beta_2 S_r |S_r| + e_r,$$

where r is a given race (i.e., state and year) and S is the snapshot. I use the absolute value of S_r, rather than the simpler square term, so that β_2 term moves the estimate further from parity (if greater than 0) or closer to parity (if less than 0), rather than up or down in absolute terms for incumbent. This β_2 term explains how the transformation differs as the snapshot indicates a wider distance between the candidates.[13] The regressions are calibrated using elections from 2000, 2004, and 2008 for the Electoral College, and 2004, 2006, 2008, and 2010 for the senatorial elections.[14] I calibrate the parameters separately for each election type and for each day before the election. Thus, I recover a unique $\alpha, \beta_1,$ and β_2 for each day before the election (T) and election type (q). The daily estimated vote share are created using those parameters: $\widehat{V_r} = \alpha_{T,q} + \beta_{1,T,q} S_r + \beta_{2,T,q} S_r |S_r|$, the alpha corrects for the anti-incumbency bias and the beta terms correct for reversion to the mean.

---

[12]The linear trend is the simplest and most transparent method to create a consistent poll average on any given day, especially in races with limited number of polls. I will show the robustness of this method in the next section.

[13]For all of parameters I use +/- 7 days of data to gain consistency, relative to the daily random variation in the Erikson and Wlezien (2008).

[14]The data is collected from: PollingReport.com, Pollster.com, and RealClearPolitics.com. Using the method from (Wlezien and Erikson, 2002) I fill in missing data, for historical data only, with the linear interpolation from the poll before and after any missing day.

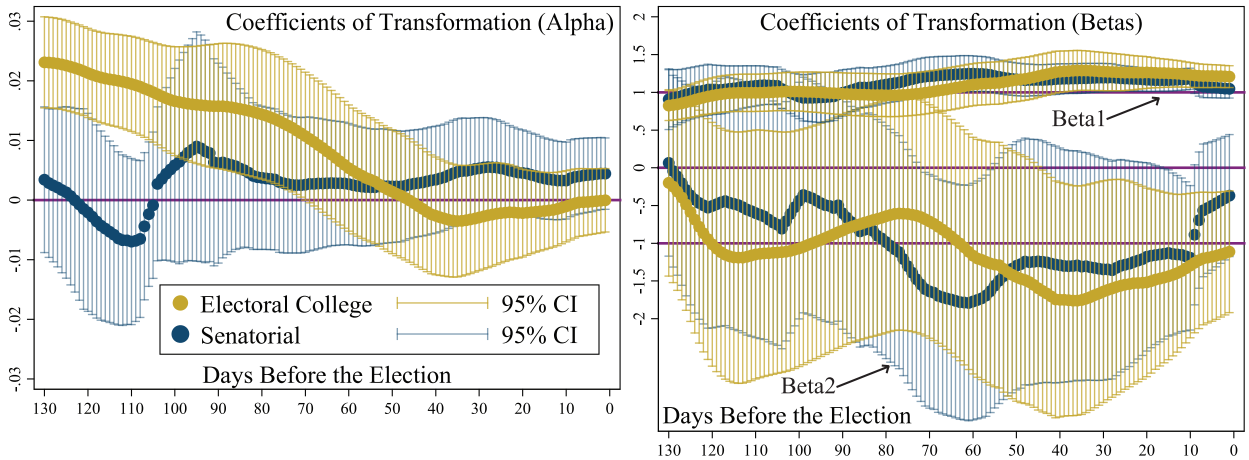

Figure 1: Alpha (left) and beta (right) for Electoral College and senatorial elections derived in equation (1). Each point plots the value of the coefficient at a given day before the elections. An alpha greater than 0 represents an anti-incumbency bias; the incumbent gains points in expectation above their snapshot. A beta1 less than 1 represents a reversion to the mean and a beta2 less than 0 means that the reversion to the mean increases as the snapshot widens. The standard errors are clustered by race (i.e., state and year).

While the Electoral College has a meaningful and statistically significant anti-incumbency bias, the senatorial elections do not; further, the results add a different perspective to the theory behind the anti-incumbency bias. Illustrated on the left side of Figure 1, a presidential candidate whose party won the state in the previous election can expect to gain one to two percentage points of the two-party vote share from his/her opponent, depending on the days before the election. This bias decreases towards zero a week or two after Labor Day. I tested coding the Electoral College in terms of the national incumbent party and the state-by-state winner from the previous election. For the model the data is in terms of the state-by-state incumbent, because it has lower forecast errors, which correlates with a more meaningful alpha term, relative to using the national incumbency. This is the first time anyone has addressed the anti-incumbency bias on the state level for the Electoral College or senatorial elections.[15]

Illustrated on the right side of Figure 1, the reversion to mean increases as the snapshot indicates a wider distance between the candidates; this is new to the literature as well. This is apparent in the large negative coefficients for the snapshot squared, while the plain snapshot's coefficient is, if it drifts away from one, actually above one. This novel result suggests that I can characterize reversion to the mean more so by a 10-point lead in the polls preceding a narrow victory, as opposed to a 2-point lead preceding a toss-up.[16]

The third step in creating a poll-based forecast is to create a probability of victory, which

[15] I also tested both Electoral College and senatorial elections in terms of party affiliation, but that provided even less accurate forecasts.

[16] While there are sizable standard errors on this coefficient, it is an extremely consistent finding, meaningful finding, and, for both types of elections, statistically significant for large periods of time.

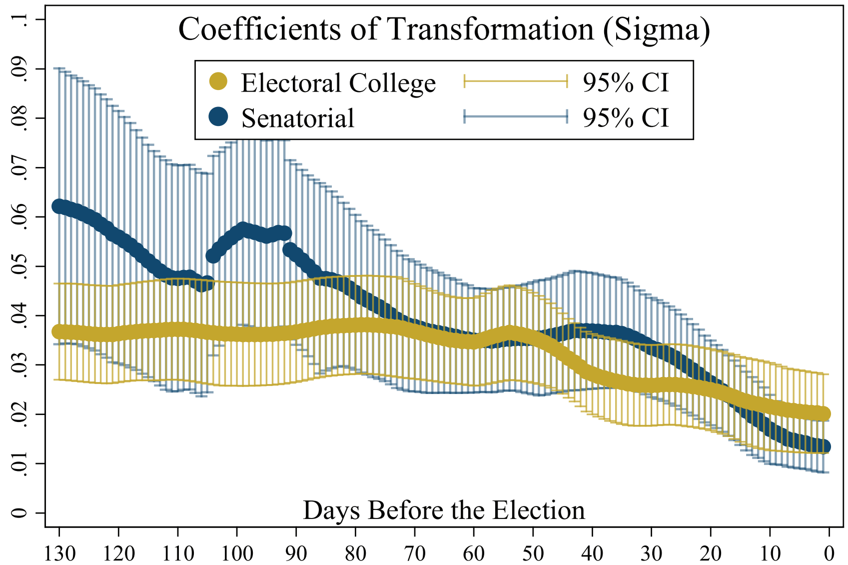

Figure 2: Sigma for Electoral College and senatorial elections derived in equation (2). Each point plots the value of the coefficient at a given day before the elections. The standard errors are clustered by race (i.e., state and year).

is the probability that the two-party vote share is greater than 50%. The method assumes that the actual vote share on Election Day is drawn from a normal distribution centered around the estimated vote share. For the same estimated vote share, the more accurate the estimation is, the tighter the distribution of true outcomes, and the greater the percentage of probable outcomes where the favored candidate has the higher amount of votes. I determine the optimal sigma ($\sigma_{T,q}$) for each day and election type by running a probit of the binary election outcomes on the expected vote share derived with the coefficients from equation (1):

$$(2) P_r = \Phi(\widehat{V_{T,q}}/\sigma_{T,q}),$$

where P_r is the probability of victory for a given race and $\widehat{V_{T,q}}$ is the estimated vote share we derived with the coefficients from equation (1). It is unsurprising that sigma gets smaller as the days before the election decrease, as shown in Figure 2. It is impressive to see how small sigma, representing the standard deviation, becomes as Election Day approaches. For most of the early days of the cycle, the sigma coefficient is larger in the senatorial versus the Electoral College, because the estimated vote shares are less accurate due to less polling and less accurate polling.[17]

The raw prediction market data translates into outcomes over a few steps. First, I take the average of the bid and ask for the stock that pays out if the Democrat wins on Election

[17]If there is a sizable correlated shock in a given cycle, then the standard deviation would be too small and the probabilities too confident. But, this method is standard, has worked historically, is relatively robust with the sample size, and I do not have enough cycles to confidently derive these coefficients out-of-sample.

Day. If the bid-ask spread is greater than five points, I take the last sale price.[18] This is the raw prediction market price. Second, I correct for the favorite-longshot bias as in (Rothschild, 2009), using the transformation suggested by (Leigh et al., 2007): $P_r = \Phi(1.64\Phi^{-1}(price))$.[19]

My datasets for examining prediction market data include just 2004 and 2008 for Electoral College and 2006, 2008, and 2010 for senatorial elections. Relatively limited prediction market data, compared to polling and fundamental data, is why I must make the combined forecast with this limited dataset. As a first step in determining the most efficient prediction market model, I duplicated the procedure in Leigh et al. (2007). I took the probit $P_r = \Phi(\beta\Phi^{-1}(price))$ and regressed it over all of the prediction market data in my datasets and recovered β=1.67; this is remarkably close to the 1.64 that Leigh et al. (2007) determined over a totally different dataset.

For both fundamental data and polling data, I create separate models for the different election types, but I do not have the identification to split prediction market prices into election type. I do not want to over-estimate the coefficients and I do not feel comfortable with just two elections cycles for the Electoral College.

But, I do want to examine two other dimensions: days before the election and the certainty of the price. The chart on the left side of Figure 3 shows how that same coefficient, β, shifts with days before the election (dbe) within the equation:

$$(3) P_r = \Phi(\beta\Phi^{-1}(price)).$$

There is no smoothing between dbe in this chart; each point is a separate regression. The coefficient is amazingly close to 1.64 the entire time, until the last few days. Second, I checked to see how the coefficient would change as the price moves away from 50. So, I ran equation (3) again for every price between 50 and 100, where I inverted prices below 50 to above 50 (i.e., 30 became 70 or all prices are in terms of the most likely candidate).[20] The chart on the right side of Figure 3 shows that the efficient coefficient is relatively stable around 1.64, regardless of the extremity of the price. The only bump is near the middle, where there is much less data and the parameter has much less impact when it is applied (i.e., even an infinite coefficient does nothing at 50). Thus, without any compelling evidence to change it, I keep 1.64 as the same coefficient to debias all prediction market data.

4. Estimation Strategy/ Results in Combining Forecasts

Comparing the three forecast types we are limited to the overlapping elections of the data which is 2004 and 2008 for Electoral College elections and 2006, 2008, and 2010 for senatorial elections; this is the data that I have for prediction markets, the narrowest dataset.

Liquidity is an issue through Election Day. From 130 days before the election I have data for all three forecast types for all Electoral College elections. The fundamental forecast

[18] Procedure is adapted from (Snowberg et al., 2006).

[19] This transformation was suggested (and estimated) prior to Rothschild (2009) in Leigh et al. (2007), using data from Presidential predication markets from 1880 to 2004.

[20] I use +/- 5 points of data to gain consistency.

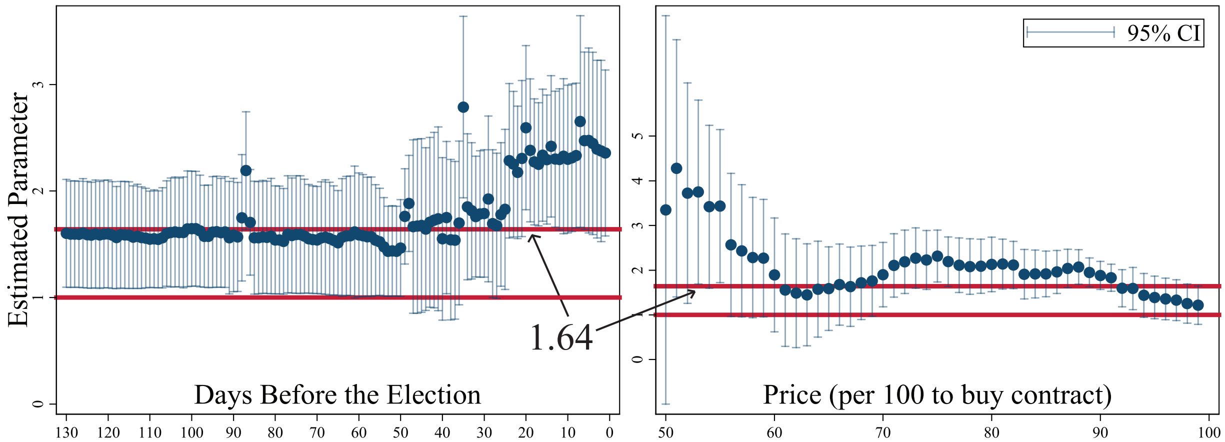

Figure 3: Sigma for Electoral College and senatorial elections in equation (3). Each point on the left panel plots the value of the coefficient at a given day before the elections. Each point on the right panel plots the value of the coefficient at a given price from 50 to 100. The standard errors are clustered by race (i.e., state and year).

provides a forecast for all elections at all times. The prediction markets can be very thin early in the cycle, but they always provide a forecast; these early forecasts can provide relatively large errors in relatively easy to predict elections if there has been little action in the market. Polls are completely absent in some senatorial elections. Out of a possible 102 senatorial elections, polling data ranges from just 41 elections 130 days before the election (dbe), to 74 at 100 dbe, to 86 at 1 dbe. Early in the cycle they are absent in a somewhat random selection of elections, while they are likely to be absent in easy-to-predict elections late in the cycle.

The three forecast types differ in their accuracy as the days before the election decrease. Figure 4 shows the mean square error of the probability of victory, relative to the final outcome (i.e., 1 for a win and 0 for a loss), for both the Electoral College and senatorial elections. No one knows the true probability of victory 50 or 100 days before an election, but the mean square error relative to the final outcome is a key rubric for determining the accuracy of forecasts over time. The accuracy of the fundamental model never moves, because it does not shift during the election.[21] The other two forecast types are very similar, with prediction markets having slightly smaller errors, with pockets of time where the difference is statistically significant. Only as the campaigns move into the main timeframe, after Labor Day, do the errors on the poll and prediction market-based forecasts drop well below the fundamental model-based forecasts. The chart is extremely similar if I use only elections where all three data types have forecasts; prediction market-based forecasts have some periods of time with statistically significantly smaller errors than poll-based forecasts early in the cycle, and very similar errors after Labor Day. Consistent with the literature, raw prediction market prices

[21]The small bump around day 100 is the addition of the West Virginia senate race in 2010, which did not exist until Senator Robert Byrds death after the cycle had already begun.

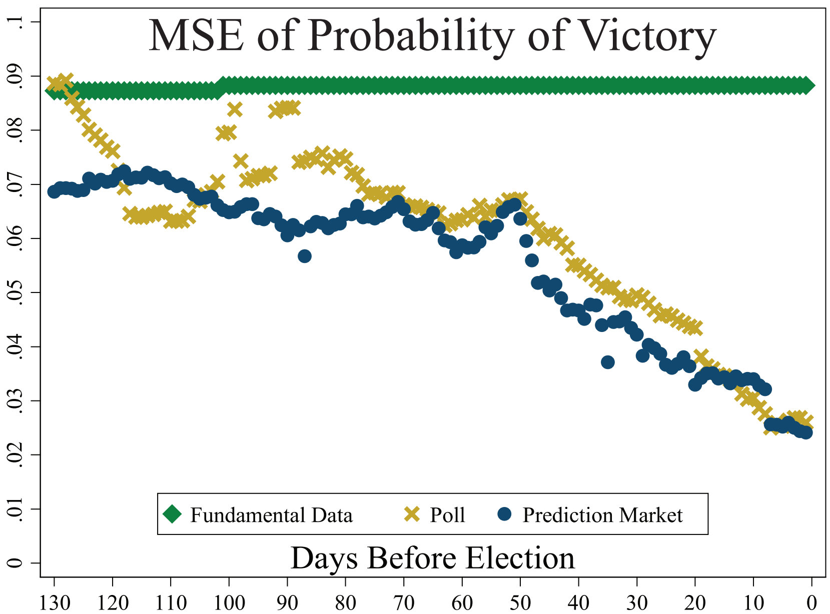

Figure 4: Accuracy of probability of victory estimates for Electoral College and senatorial elections by fundamental data, voter intention poll, and prediction markets-based forecasts. There are 202 observations per dbe for the combined and prediction market-based forecasts. If poll data does not exist, prediction market data is takes its place in combined forecast. Polls-based forecasts vary from 141 to 186 observations per dbe.

have regularly higher errors than properly debiased prediction market-based forecast and not debiasing polls leaves raw polls with substantially larger errors than properly debiased poll-based forecasts.

I can combine forecasts, to create the combined forecast, in many different ways, but I am even more concerned here than in the individual forecasts to ensure that I avoid over-estimating the coefficients. Again, with only two presidential cycles, I avoid any attempt to separate the parameters by election type. Accuracy by forecast should shift with time, which is highly correlated with variations in the quantity of polls available and liquidity and prediction markets. In the interest of simplicity I do not allow the aggregation to vary for different forecasts types within day. Thus, I focus exclusively on the days before the election, which Figure 4 shows should be a major factor.

I combine the forecasts for probability of victory very directly with a probit of the inverse normal of their probabilities:

$$(4) P_r = \Phi((\Sigma\beta + \Sigma\gamma dbe)(\Phi^{-1}(P_{r,F}) + \Phi^{-1}(P_{r,Poll}) + \Phi^{-1}(P_{r,PM})),$$

where I allow the parameters to shift linearly by the dbe.[22]

[22] Allowing the coefficients to vary daily was over-estimated. By regressing across time I introduce correlated error, so I cluster by race. Since individual races are not always independent this should be taken as a lower bound of error; but, that should not affect the point-estimates of the coefficients.

This produces a very clean result; when everything is added together, the coefficients vary from approximately equal weighting of each forecast on 130 dbe, to averaging just the poll and prediction market-based forecasts on Election Day. By the definition of the model, the weight on the fundamental data moves in a linear manner. The derived coefficients are in Table 1.

To keep everything simple and to avoid over-estimating the model, I simplify a few things when I use this model to create out-of-sample real-time forecasts. First, I drop insignificant variables, which leaves just β for polls and prediction markets, and γ for fundamental data. Second, I recalibrate the coefficients so that they sum to 1 at any given dbe; I do this because I have no reason to believe that they should go beyond 1 and I want to be conservative until I have more election cycles. Finally, as it is already very close, I round the coefficients to exactly 0.333 each at 130 days before the election and allow them to move linearly until they are 0.5 each for the polling and prediction market-based forecasts, and 0 for fundamental-based forecasts, on Election Day.

Variable	Coefficients for Probability of Victory
Fundamental Data	0.229 (0.125)
Polling	0.645* (0.247)
Prediction Market	0.726* (0.210)
Fundamental Data*dbe	0.003* (0.001)
Polling*dbe	-0.001 (0.003)
Prediction Market*dbe	-0.003 (0.003)

Table 1: The coefficients for combining the three forecast types into a combined probability of victory, equation (4): $P_r = \Phi((\Sigma\beta + \Sigma\gamma dbe)(\Phi^{-1}(P_{r,F}) + \Phi^{-1}(P_{r,Poll}) + \Phi^{-1}(P_{r,PM}))$. The standard errors are clustered by race (i.e., state and year). * denotes significance at the 5% level.

These results so far are within-sample, so, despite the simplifications of the coefficients, it is almost by definition that the combined forecast is more accurate than the individual forecasts.[23] Illustrated in Figure 5 , it is not surprising that combining the three forecasts into the combined forecast provides the most potential benefit early in the cycle where the fundamental data is heavily involved and there is wide variation in the accuracy of the forecasts. Late in the cycle the poll and prediction market forecasts converge, so, by definition, they converge with a combined forecast which is mainly a combination of the two forecasts. However, the combined forecast generally performs better than each individual forecast for almost any given dbe. The combined forecast is equally dominant if I only use elections where we have all three data types.

Researchers should judge forecasts on calibration as well as the size of the errors. Calibration is unique to probability of victory; it measures how often an event occurs relative to the forecasted probability. For example, if a properly calibrated forecast declares 100 events about 75% likely to occur, then seventy-five of the events should occur. The goal

[23]I state almost by definition, because I simplified the parameters from the regression results, which will raise the mean square error of the forecast.

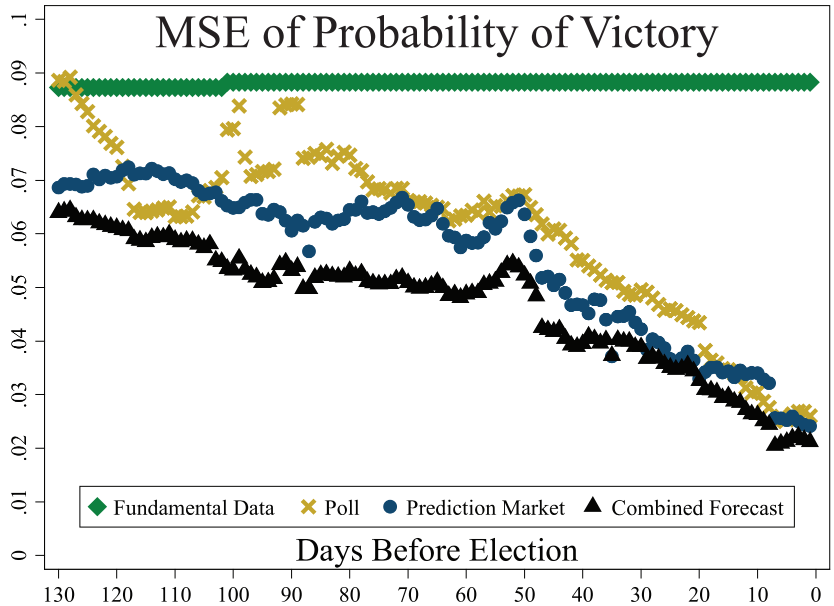

Figure 5: Accuracy of probability of victory estimates for Electoral College and senatorial elections by fundamental data, voter intention poll, and prediction markets-based forecasts, along with combined forecast.There are 202 observations per dbe for the combined and prediction market-based forecasts. If poll data does not exist, prediction market data is takes its place in combined forecast. Polls-based forecasts vary from 141 to 186 observations per dbe.

of a calibration is to see how well the forecast discriminates certainty; this metric actually rewards a well calibrated forecast of toss-up, while the error would be large regardless of the outcome. Figure 6 shows the relationship between the probability of the leading candidate winning (so all forecasts are between 50% and 100%) and the percent of elections won by the leading candidate. Every probability is rounded to the nearest 5% mark and the circles sizes correlate with the quantity of observations in that bucket. A well-calibrated forecast is close to the 45 degree line shown in purple; the combined forecast is generally close to the 45 degree line as you move from the lowest probabilities, which occur with low frequently, to the highest probabilities, which occur with high frequency.

I do not want this paper to dwell too much on a single year, as the correlation between the outcomes somewhat diminishes the explanatory power of even 83 different outcomes.[24] Yet, the combined forecast does well in predicting the 2012 election out-of-sample; Figure 7 shows the errors every 4 hours for the last 130 days of the election in 2012. Again, I created the model for the combined forecast prior to the 2012 election; a publicly available website published the combined forecast during the election and updated it every few minutes.[25] Unlike the within-sample years it was not dominant at every point in the cycle, but it was

[24]We drop DC's Electoral College election.

[25]The combined forecast had well over 1,000,000 pageviews from October 1, 2012 through Election Day 2012.

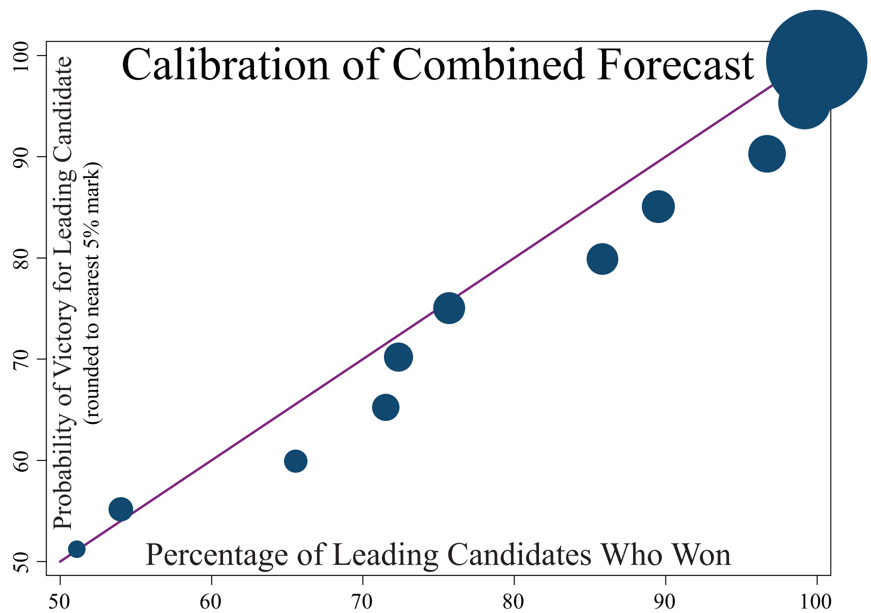

Figure 6: Calibration of probability of victory estimates for Electoral College and senatorial elections by the combined forecast. There are 202 observations per dbe for the combined forecasts and this chart is aggregated over the final 130 dbe. If poll data does not exist, prediction market data is takes its place. Polls-based forecasts vary from 141 to 186 observations per dbe. The size of the circles correlates with the quantity of observations in that bucket. The bucket sizes range from 315 near 50% to 16,303 near 100%. The quantity increases monotonically.

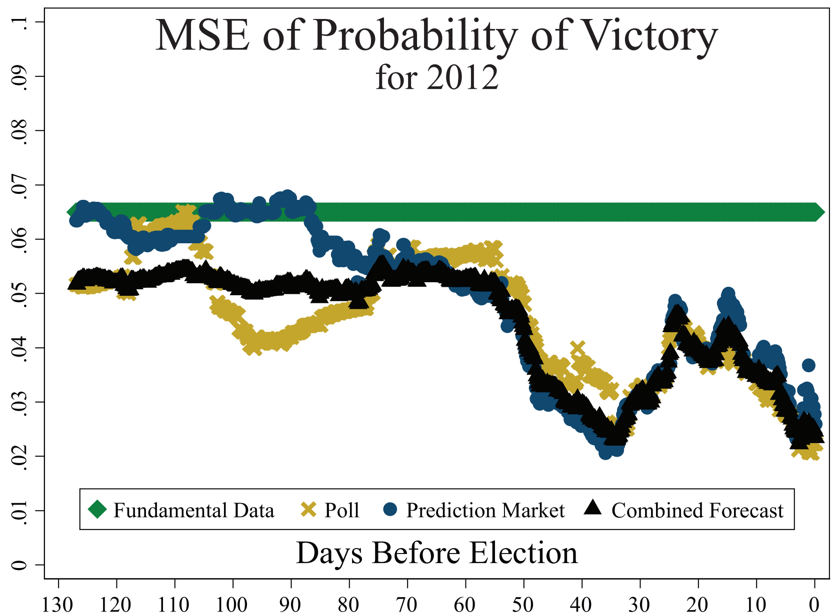

Figure 7: Accuracy of probability of victory estimates for Electoral College and senatorial elections by fundamental data, voter intention poll, and prediction markets-based forecasts, along with combined forecast for 2012. There are 83 observations per dbe for all forecasts, at all dbe.

the most consistent forecast. For a span of about 30 days early in the cycle when poll-based forecasts had a lower error than prediction market-based forecasts, the combined forecast was either below or near the poll-based forecast. Towards the end of the summer until last the month of the campaign, a span of about 45 days when prediction market-based forecast had a lower error than polls, the combined forecast again held closely to the lowest errors. At any given moment from 130 before the election to Election Day in 2012 the combined forecast is likely to have a lower error than either the poll-based or prediction market-based forecast.[26]

5. Estimation Strategy/ Results in Creating Separate Forecasts: Expected Vote Share

Both fundamental data (in Hummel and Rothschild (2013)) and polling (within the last section in equation (1)) are already translated into estimated vote share; prediction market data translates into estimated vote share by regressing the inverse of the price on the vote

[26]The combined forecast and prediction market-based forecast use the average of the raw prices from Betfair and Intrade, when both markets were available. The combined forecast and poll-based forecast are tested for both the poll aggregation method described in this paper and the more complicated aggregation method utilized by Simon Jackman for Pollster.com, where both are then translated into a forecast with the same method; while the method used in this paper has smaller error on average, the difference between the two aggregation methods is not significant to the findings.

share and predicting the value:

$$(5) V_r = \beta \Phi^{-1}(price).$$

There should be no meaningful distinction in the probabilities of a big certain wins and a small certain win (i.e., if Candidate A is estimated to receive 51 percent of the vote in her election and Candidate B is estimated to receive 75 percent of the vote in a different election, but both are certain to win their respective elections, both of their prediction market prices will approach \$1.00 for \$1.00 contracts, providing no identification for their estimated vote shares).

The combined forecast of expected vote share is determined in the same manner as probability of victory. I use the following regression:

$$(6) V_r = (\Sigma\beta + \Sigma\gamma dbe)(\widehat{V_{r,F}} + \widehat{V_{r,Poll}} + \widehat{V_{r,PM}}),$$

where I allow the coefficients to shift linearly by the dbe. All coefficients, except the β for prediction markets, are highly significant. The derived coefficients are Table 2. The poll-based forecast is weighted the most for the entire length of the campaign, starting with just a little more weight 130 days before the election to over 80% of the weight by Election Day. The absolute error of the resulting combined forecast is compared with the three individual forecasts in Figure 8. Not surprisingly, the forecast converges towards the poll forecast as the Election Day approaches.[27]

Variable	Coefficients for Estimated Vote Share
Fundamental Data	0.173* (0.036)
Polling	0.886* (0.072)
Prediction Market	0.001 (0.051)
Fundamental Data*dbe	0.001*(0.000)
Polling*dbe	-0.004* (0.001)
Prediction Market*dbe	0.002* (0.001)

Table 2: The coefficients for combining the three forecast types into a combined estimated vote share, equation (6): $V_r = (\Sigma\beta + \Sigma\gamma dbe)(\widehat{V_{r,F}} + \widehat{V_{r,Poll}} + \widehat{V_{r,PM}})$. The standard errors are clustered by race (i.e., state and year). * denotes significance at the 5% level.

6. Conclusion

This paper combines three forecasts based on polling data, fundamental data, and prediction market data. This combined forecast accounts for shifting levels of information, by allowing the parameters to adjust with the relative value of the separate data types over time.

[27]This result is validates and expands (Erikson and Wlezien, 2012) which determines that prediction markets do not provide information in estimating the nation popular vote. Of course, a central thesis of this paper is that that is not the correct outcome variable for relevant election forecasts.

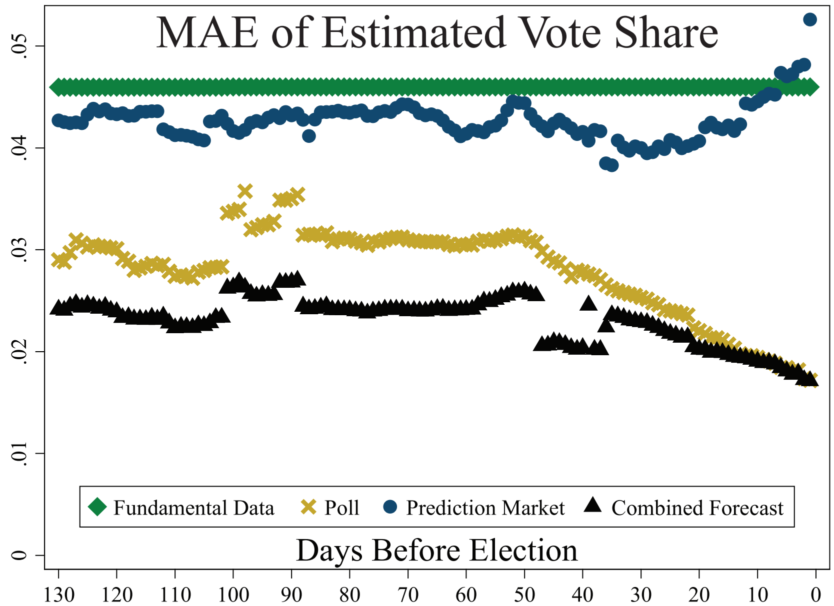

Figure 8: Accuracy of estimated vote share for Electoral College and senatorial elections by fundamental data, voter intention poll, and prediction markets-based forecasts, along with combined forecast. There are 202 observations per dbe for the combined and prediction market-based forecasts. If poll data does not exist, fundamental data is takes its place in combined forecast. Polls-based forecasts vary from 141 to 186 observations per dbe.

The paper shows that all three data types are meaningful and significant in an efficient forecast of state-by-state elections; to create efficient forecasts, it is crucial that researchers are agnostic toward different data types. The benefit of combining is highest earlier in the cycle, as polls and prediction markets converge towards Election Day and, thus, the combined forecast becomes very close to either of the two forecasts.

A secondary finding of the paper is to provide new insights into the translation of raw polling and prediction market data into forecasts. For example the anti-incumbency bias for polls does not extend to senatorial elections. This is likely due to the longer length between elections that allows for more change and the massive jumps between parties that can occur with periodic retirements. In another example, the favorite-longshot bias for prediction markets does not vary significantly by time. In theory there are more liquidity concerns the longer the contract is held, as the opportunity costs increase and the liquidity in the market is lower, which should increase the favorite-longshot bias with days before the election. But, this does not seem to impact anything within a 130 day window.

Some researchers may be surprised that fundamental data provides anything not found in either polls or prediction markets, but fundamental data provides two things: stability and liquidity. First, fundamental forecasts provide stability, when prediction markets or polls may be chasing short-term fluctuations in sentiment. This was certainly in the case in the 2012 presidential election when Mitt Romney's bump after the first debate ultimately dissipated. Second, both polls and prediction markets are not that well formed 130 days before the election. Only after Labor Day are the polls and prediction markets both fully liquid.[28]

Some researchers may be surprised that prediction markets add any information beyond polls, but there are three types of data to consider. First, prediction markets can incorporate the impact of major events that polls take several days to acquire (e.g., the release of a secret video). Second, prediction markets can incorporate the impact of events that have not yet occurred, but users know will impact the electorate (e.g., the likely impact of a debate). Finally, prediction markets can aggregate idiosyncratic information about the election from self-selected users with high levels of information that poll respondents are not able to, or do not have, to incorporate into their poll responses (e.g., major differences in get-out-the-vote efforts).

The value of the regularly updating nature of the forecast, which can only be provided by prediction market data, was seen in 2012. Forecasts with only fundamental and/or polling data would not update around the conventions or debates. With only Labor Day weekend between the Republican and Democratic debates, it was difficult for most polling companies to field polls that show the impact of the conventions separately. Similarly, while observers could speculate on how the polls would move in the days following Romney's triumph over Obama in the first debate, the combined forecast moved during, and in the immediate aftermath, of the debate.

Finally, some researchers may be surprised that polls add any information beyond prediction markets. In theory prediction markets should already include all of the latest polling information. It is possible that some of the prediction market data suffers from enough liquidity issues, at various points in the cycle, so that the benefit of aggregation with other data types is to correct those issues. Further, prediction markets are prone to small price shifts from users whose incentives are not to maximize their return with the market, but to hedge other investments or to influence the campaign (Rothschild and Sethi, 2013). This paper does not challenge the efficient market principle in general, but shows empirically that polls are necessary to fill holes in the information powering a forecast that runs continuously.

The combined forecast is a practical forecast; a publicly available website published the forecasts from this model during the election cycle. Publishing the forecast during the campaign not only ensured that 2012 was cleanly out-of-sample, but it forced me to consider data issues that can be ignored if the forecast is run ex-post. For example, some states do not hold their senatorial primaries until the second week in September. Since senatorial polling is candidate-to-candidate, this makes using polling data early in the cycle difficult, as I had to weigh forecasts by the likelihood of victory in the primary, but it is no problem after the election when a forecaster can determine with certainty which matchup occurred.

More refined combinations of data are certainly possible, but I am wary of over-estimating my models. There are not enough elections and not enough independence within elections

[28]Fundamental forecasts are still interesting in October to explain correlations between the world the election (e.g., people want to know the expected impact of new economic indicators). But, they should not be confused with an updated and accurate forecast.

to identify too intricate models. As it is, my standard errors could be slightly higher due to correlations between outcomes. That is why I round my coefficients for the combined forecast and try to be conservative as possible with these results. While the question of standard errors is not very germane to my generalized results, it should serve as warning to future research that attempts to estimate their coefficients too tightly.

The model in this paper is designed to be easily duplicated in real-time for future elections. Future testing will consider the benefits of allowing the coefficients for combining forecasts to vary by race, based on metrics of certainty. This would require additional data that I do not have historically, but, after collecting that data, I would have to weigh any possible benefits of accuracy with offsetting costs of data collections, running the model in real-time, and possible over-estimation. The central question would be whether or not there is enough identification to accurately model both shifting weights by day, which is currently in the model, and by certainty of the data types within day.

This paper takes steps to clarify the objectives of creating forecasts to provide the most efficient forecast for relevant stakeholders. This includes defining the relevant outcome variable as probability of victory on a state-by-state level, ensuring that the forecasts update continuously during the course of the main election cycle, and testing calibration along with traditional measures of error.

I have framed this paper on making the most relevant, timely, and accurate forecast, but that ignores a key fourth component: cost efficiency. Fundamental forecasts are extremely costly in that researchers need to derive new models for every outcome. Standard polling is becoming increasingly expensive as response rates plummet. Polling exists in senatorial races, but is extremely sparse. Prediction markets can expand with no marginal cost, but can have trouble getting liquidity for some events. Overall, prediction markets are the most cost efficient, but different situations will vary this relationship. As we move into a digital age, with more data and way to contact people, we should be mindful that more cost efficient forecasts will help us answer more questions in more domains. That will add efficiency to our decisions and allow researchers to answer new questions.

Appendix A. Coefficients for Fundamental Model

The coefficients for expected vote share from Hummel and Rothschild (2013) are in the text; here are the coefficients for probability of victory which should be input into a probit. Electoral College model has the following coefficients for Democratic likelihood of victory: 0.15 on (presidential approval - 42)*incumbency, -0.64 on two or more terms*incumbency, 0.21 on state vote four years ago - national vote, 0.07 on state vote four years ago - national vote, 0.15 on change in state income from 9th to 13th quarter of term, -0.005 on sum of ACU rankings for senators - average sum of ACU senators, 0.05 on ..., 1.69 on home state if less than 10 million in population, and -0.89 as a constant. The senatorial model has the following coefficients for Democratic likelihood of victory: 0.02 on (presidential approval - 50)*presidential incumbency, 1.34 on incumbency, -0.51 on midterm*presidential incumbency, 0.06 on last presidential vote - national vote, 0.03 on state vote six years ago - national vote, 0.04 on change in state income from 9th to 13th quarter of term*presidential incumbency, between 0.55 and 1.05 on previous job of senatorial candidates, and 0.15 as a constant. There are less variables in the probability of victory versus estimated vote share, because some variable are significant for size of victory, but not the binary outcome.

Abramowitz, A. I., 2004. When good forecasts go bad: The time-for-change model and the 2004 presidential election. PS-WASHINGTON- 37 (4), 745–746.

Abramowitz, A. I., 2008. Forecasting the 2008 presidential election with the time-for-change model. PS: Political Science and Politics 41 (4), 691–95.

Armstrong, J. S., Green, K. C., Jones, R. J., Wright, M. J., 2010. Predicting elections from politicians faces. International Journal of Public Opinion Research 22 (4), 511–522.

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

Campbell, J. E., 2008. The American campaign: US presidential campaigns and the national vote. Vol. 6. Texas A&M University Press.

Campbell, J. E., 2012. Forecasting the 2012 american national elections. PS: Political Science & Politics 45 (04), 610–613.

Clemen, R. T., 1989. Combining forecasts: A review and annotated bibliography. International Journal of Forecasting 5 (4), 559–583.

Erikson, R. S., Wlezien, C., 2008a. Are political markets really superior to polls as election predictors? Public Opinion Quarterly 72 (2), 190–215.

Erikson, R. S., Wlezien, C., 2008b. Leading economic indicators, the polls, and the presidential vote. PS: Political Science and Politics 41 (4), 703–707.

Erikson, R. S., Wlezien, C., 2012. Markets vs. polls as election predictors: An historical assessment. Electoral Studies 31 (3), 532–539.

Graefe, A., Armstrong, J. S., Jones Jr, R. J., Cuzán, A. G., 2012. Combining forecasts: An application to elections. In: 2011 APSA Annual Meeting Paper.

Hummel, P., Rothschild, D., 2013. Fundamental models for forecasting elections. ResearchDMR. com/HummelRothschild_FundamentalModel.

Jackman, S., 2005. Pooling the polls over an election campaign. Australian Journal of Political Science 40 (4), 499–517.

Klarner, C., 2008. Forecasting the 2008 us house, senate and presidential elections at the district and state level. PS: Political Science & Politics 41 (04), 723–728.

Klarner, C. E., 2012. State-level forecasts of the 2012 federal and gubernatorial elections. PS: Political Science & Politics 45 (04), 655–662.

Leigh, A., Wolfers, J., Zitzewitz, E., 2007. Is there a favorite-longshot bias in election markets. In: Preliminary version presented at the 2007 UC Riverside Conference on Prediction Markets, Riverside, CA.

Lewis-Beck, M. S., Rice, T. W., 1992. Forecasting elections. CQ Press.

Linzer, D. A., 2013. Dynamic bayesian forecasting of presidential elections in the states. Journal of the American Statistical Association 108 (501), 124–134.

Lock, K., Gelman, A., 2010. Bayesian combination of state polls and election forecasts. Political Analysis 18 (3), 337–348.

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

Montgomery, J. M., Hollenbach, F. M., Ward, M. D., 2012. Improving predictions using ensemble bayesian model averaging. Political Analysis 20 (3), 271–291.

Pickup, M., Johnston, R., 2005. Measurement error and house bias in 2004 presidential campaign polls. In: Annual Meeting of the American Political Science Association, Washington, DC, September. pp. 2–5.

Rhode, P. W., Strumpf, K. S., 2004. Historical presidential betting markets. The Journal of Economic Perspectives 18 (2), 127–141.

Rothschild, D., 2009. Forecasting elections comparing prediction markets, polls, and their biases. Public Opinion Quarterly 73 (5), 895–916.

Rothschild, D. M., Sethi, R., 2013. Trading strategies and market microstructure: Evidence from a prediction market. Available at SSRN 2322420.

Snowberg, E., Wolfers, J., 2010. Explaining the favorite-longshot bias: Is it risk-love or misperceptions? Tech. rep., National Bureau of Economic Research.

Snowberg, E., Wolfers, J., Zitzewitz, E., 2006. Party influence in congress and the economy. Tech. rep., National Bureau of Economic Research.

Wlezien, C., Erikson, R. S., 2002. The timeline of presidential election campaigns. The Journal of Politics 64 (04), 969–993.

Wolfers, J., Zitzewitz, E., 2006. Interpreting prediction market prices as probabilities. Tech. rep., National Bureau of Economic Research.