Forecasting Elections: Voter Intentions versus Expectations*

David Rothschild
Microsoft Research
and Applied Statistics Center, Columbia

David@ReseachDMR.com
www.ResearchDMR.com

Justin Wolfers
Dept of Economics and Ford School of Public Policy,
University of Michigan
Brookings, CEPR, CESifo, IZA and NBER
jwolfers@umich.edu
www.nber.org/~jwolfers

Abstract

Most pollsters base their election projections off questions of voter intentions, which ask "If the election were held today, who would you vote for?" By contrast, we probe the value of questions probing voters' expectations, which typically ask: "Regardless of who you plan to vote for, who do you think will win the upcoming election?" We demonstrate that polls of voter expectations consistently yield more accurate forecasts than polls of voter intentions. A small-scale structural model reveals that this is because we are polling from a broader information set, and voters respond as if they had polled twenty of their friends. This model also provides a rational interpretation for why respondents' forecasts are correlated with their expectations. We also show that we can use expectations polls to extract accurate election forecasts even from extremely skewed samples.

This draft: January 23, 2013

Keywords: Polling, information aggregation, belief heterogeneity

JEL codes: C53, D03, D8

* The authors would like to thank Stephen Coate, Alex Gelber, Andrew Gelman, Sunshine Hillygus, Mat McCubbins, Marc Meredith and Frank Newport, for useful discussions, and seminar audiences at AAPOR, Berkeley, Brookings, CalTech, Columbia, Cornell, the Conference on Empirical Legal Studies, the Congressional Budget Office, the Council of Economic Advisers, Harvard, Johns Hopkins, Maryland, Michigan, MIT, the NBER Summer Institute, University of Pennsylvania's political science department, Princeton, UCLA, UCSD, USC, and Wharton for comments.

I. Introduction

Since the advent of scientific polling in the 1930s, political pollsters have asked people whom they intend to vote for; occasionally, they have also asked who they think will win. Our task in this paper is long overdue: we ask which of these questions yields more accurate forecasts. That is, we evaluate the predictive power of the questions probing voters' *intentions* with questions probing their *expectations*. Judging by the attention paid by pollsters, the press, and campaigns, the conventional wisdom appears to be that polls of voters' intentions are more accurate than polls of their expectations.

Yet there are good reasons to believe that asking about expectations yields more greater insight. Survey respondents may possess much more information about the upcoming political race than that probed by the voting intention question. At a minimum, they know their own current voting intention, so the information set feeding into their expectations will be at least as rich as that captured by the voting intention question. Beyond this, they may also have information about the current voting intentions—both the preferred candidate and probability of voting—of their friends and family. So too, they have some sense of the likelihood that today's expressed intention will be changed before it ultimately becomes an election-day vote. Our research is motivated by idea that the richer information embedded in these expectations data may yield more accurate forecasts.

We find robust evidence that polls probing voters' expectations yield more accurate predictions of election outcomes than the usual questions asking about who they intend to vote for. By comparing the performance of these two questions only when they are asked of the exact same people in exactly the same survey, we effectively difference out the influence of all other factors. Our primary dataset consists of all the state-level electoral presidential college races from 1952 to 2008, where both the intention and expectation question are asked. In the 77 cases in which the intention and expectation question predict different candidates, the expectation question picks the winner 60 times, while the intention question only picked the winner 17 times. That is, 78% of the time that these two approaches disagree, the expectation data was correct. We can also assess the relative accuracy of the two methods by assessing the extent to which each can be informative in forecasting the final vote share; we find that relying on voters' expectations rather than their intentions yield substantial and statistically significant increases in forecasting accuracy. An optimally-weighted average puts over 90% weight on the expectations-based forecasts. Once one knows the results of a poll of voters expectations, there is very little additional information left in the usual polls of voting intentions. Our findings remain robust to correcting for an array of known biases in voter intentions data.

The better performance of forecasts based on asking voters about their expectations rather than their intentions, varies somewhat, depending on the specific context. The expectations question performs particularly well when: voters are embedded in heterogeneous (and thus, informative) social networks; when they don't rely too much on common information; when small samples are involved (when the extra information elicited by asking about intentions counters the large sampling error in polls of intentions); and at a point in the electoral cycle when voters are sufficiently engaged as to know what their friends and family are thinking.

Our findings also speak to several existing strands of research within election forecasting. A literature has emerged documenting that prediction markets tend to yield more accurate forecasts than polls (Wolfers and Zitzewitz, 2004; Berg, Nelson and Rietz, 2008). More recently, Rothschild (2009) has updated these findings in light of the 2008 Presidential and Senate races, showing that forecasts based on prediction markets yielded systematically more accurate forecasts of the likelihood of Obama winning each state than did the forecasts based on aggregated intention polls compiled by Nate Silver for the website FiveThirtyEight.com. One hypothesis for this superior performance is that because prediction markets ask traders to bet on outcomes, they effectively ask a different question, eliciting the expectations rather than intentions of participants. If correct, this suggests that much of the accuracy of prediction markets could be obtained simply by polling voters on their expectations, rather than intentions.

These results also speak to the possibility of producing useful forecasts from non-representative samples (Robinson, 1937), an issue of renewed significance in the era of expensive-to-reach cellphones and cheap online survey panels. Surveys of voting intentions depend critically on being able to poll representative cross-sections of the electorate. By contrast, we find that surveys of voter expectations can still be quite accurate, even when drawn from non-representative samples. The logic of this claim comes from the difference between asking about expectations, which may not systematically differ across demographic groups, and asking about intentions, which clearly do. Again, the connection to prediction markets is useful, as Berg and Rietz (2006) show that prediction markets have yielded accurate forecasts, despite drawing from an unrepresentative pool of overwhelmingly white, male, highly educated, high income, self-selected traders.

While questions probing voters' expectations have been virtually ignored by political forecasters, they have received some interest from psychologists. In particular, Granberg and Brent (1983) document wishful thinking, in which people's expectation about the likely outcome is positively correlated with what they want to happen. Thus, people who intend to vote Republican are also more likely to predict a Republican victory. This same correlation is also consistent with voters preferring the candidate they think will win, as in bandwagon effects, or gaining utility from being optimistic. We re-interpret this correlation through a rational lens, in which the respondents know their own voting intention with certainty and have knowledge about the voting intentions of their friends and family.

Our alternative approach to political forecasting also provides a new narrative of the ebb and flow of campaigns, which should inform ongoing political science research about which events really matter. For instance, through the 2004 campaign, polls of voter intentions suggested a volatile electorate as George W. Bush and John Kerry swapped the lead several times. By contrast, polls of voters' expectations consistently showed the Bush was expected to win re-election. Likewise in 2008, despite volatility in the polls of voters' intentions, Obama was expected to win in all of the last 17 expectations polls taken over the final months of the campaign. And in the 2012 Republican primary, polls of voters intentions at different points showed Mitt Romney trailing Donald Trump, then Rick Perry, then Herman Cain, then Newt Gingrich and then Rick Santorum, while polls of expectations showed him consistently as the likely winner.

We believe that our findings provide tantalizing hints that similar methods could be useful in other forecasting domains. Market researchers ask variants of the voter intention question in an array of contexts, asking questions that elicit your preference for one product, over another. Likewise, indices of consumer confidence are partly based on the stated purchasing intentions of consumers, rather than their expectations about the purchase conditions for their community. The same insight that motivated our study—that people also have information on the plans of others—is also likely relevant in these other contexts. Thus, it seems plausible that survey research in many other domains may also benefit from paying greater attention to people's expectations than to their intentions.

The rest of this paper proceeds as follows, In Section II, we describe our first cut of the data, illustrating the relative success of the two approaches to predicting the winner of elections. In Sections III and IV, we focus on evaluating their respective forecasts of the two-party vote share. Initially, in Section III we provide what we call naïve forecasts, which follow current practice by major pollsters; in Section IV we product statistically efficient forecasts, taking account of the insights of sophisticated modern political scientists. Section V provides out-of-sample forecasts based on the 2008 election. Section VI extends the assessment to a secondary data source which required substantial archival research to compile. In Section VII, we provide a small structural model which helps explain the higher degree of accuracy obtained from surveys of voter expectations. Section VIII characterizes the type of information that is reflected in voters' expectation, arguing that it is largely idiosyncratic, rather than the sort of common information that might come from the mass media. Section IX assesses why it is that people's expectations are correlated with their intentions. Section VI uses this model to show how we can obtain surprisingly accurate expectation-based forecasts with non-representative samples. We then conclude. To be clear about the structure of the argument: In the first part of the paper (through section IV) we simply present two alternative forecasting technologies and evaluate them, showing that expectations-based forecasts outperform those based on traditional intentions-based polls. We present these data without taking a strong position on why. But then in later sections we turn to trying to assess what explains this better performance. Because this assessment is model-based, our explanations are necessarily based on auxiliary assumptions (which we spell out).

Right now, we begin with our simplest and most transparent comparison of the forecasting ability of our two competing approaches.

II. Forecasting the Election Winner

Our primary dataset consists of the American National Election Studies (ANES) cumulative data file for 1948-2008, which is the only research dataset that has systematically asked about voter expectations.

Forecasting the winner of the national election

In particular, we are interested in responses to two questions:

Voter Intention: *Who do you think you will vote for in the election for President?*

Voter Expectation: *Who do you think will be elected President in November?*

These questions are typically asked one month prior to the election. Throughout this paper, we treat elections as two-party races, and so discard those responses either intending to vote for a third party, or which expect a third party candidate to win. In order to keep the sample sizes comparable, we only keep respondents with valid responses to both the intention and expectation questions and our analysis of polling data uses the provided weights. When we describe the "winner" of an election, we are thinking about the outcome that most interests forecasters, which is who takes office (and so we describe George W. Bush as the winner of the 2000 election, despite his losing the popular vote).

At the national level, both questions have been asked since 1952, and to give a sense of the basic patterns, we summarize these data in Table 1.

Table 1: Forecasting the Winner of the Presidential Races

Year	Race	%Expect the winner	%Intended to vote for winner	%Reported voting for winner	Actual result: %Voting for winner	Sample size
1952	Eisenhower beat Stevenson	56.0%	56.0%	58.6%	55.4%	1,135
1956	Eisenhower beat Stevenson	76.4%	59.2%	60.6%	57.8%	1,161
1960	Kennedy beat Nixon	<u>45.0%</u>	<u>45.0%</u>	<u>48.4%</u>	50.1%	716
1964	Johnson beat Goldwater	91.0%	74.1%	71.3%	61.3%	1,087
1968	Nixon beat Humphrey	71.2%	56.0%	55.5%	50.4%	844
1972	Nixon beat McGovern	92.5%	69.7%	68.7%	61.8%	1,800
1976	Carter beat Ford	52.6%	51.4%	50.3%	51.1%	1,320
1980	Reagan beat Carter	<u>46.3%</u>	<u>49.5%</u>	56.5%	55.3%	870
1984	Reagan beat Mondale	87.9%	59.8%	59.9%	59.2%	1,582
1988	GHW Bush beat Dukakis	72.3%	53.1%	55.3%	53.9%	1,343
1992	Clinton beat GHW Bush	65.2%	60.8%	61.5%	53.5%	1,541
1996	Clinton beat Dole	89.6%	63.8%	60.1%	54.7%	1,274
2000	GW Bush beat Gore	<u>47.4%</u>	<u>45.7%</u>	<u>47.0%</u>	49.7%	1,245
2004	GW Bush beat Kerry	67.9%	<u>49.2%</u>	51.6%	51.2%	921
2008	Obama beat McCain	65.7%	56.6%	56.5%	53.7%	1,632
	Simple Average:	**68.5%**	**56.7%**	**57.4%**	**54.6%**	**18,471**

<u>*Notes:*</u> Table summarizes authors' calculations, based on data from the American National Election Studies, 1952-2008. Reported proportions are shares of votes, intentions or expectations on a two-party basis. Sample restricted to respondents whose responses to both the expectation and intention questions listed the two major candidates; *n*=18,471. Underlined entries highlight incorrect forecasts.

Each method can be used to generate a forecast of the most likely winner, and so we begin by assessing how often the majority response to each question correctly picked the winner. The first column of data in Table 1 shows that the winning Presidential candidate was expected to win by a majority of respondents in 12 of the 15 elections, missing Kennedy's narrow victory in 1960, Reagan's election in 1980, and G.W. Bush's controversial win in 2000. The more standard voter intention question performed slightly worse, correctly picking the winning candidate in one fewer election. The only election in which the two approaches pointed to different candidates was 2004, in which a majority of respondents correctly expected that Bush would win, while a majority intended to vote for Kerry. So far we have been analyzing data from the pre-election interviews. In the third column we summarize data from post-election interviews which also ask which candidate each respondent ultimately voted for. The data in this column reveal the influence of sampling error, as a majority of the people sampled in 1960 and 2000 ultimately did vote for the losing candidate.

The last line of this table summarizes, showing that on average, 68.5% of all voters correctly expected the winner of the Presidential election, while 56.7% intended to vote for the winner. These averages give a hint as to the better performance of expectation-based forecasts. Taken literally, they say that if we forecasted election outcomes based on a random sample of one person, asking about voter expectations would predict the winner 68.5% of the time, compared to 56.7%, when asking about voter intentions. More generally, in small polls, sampling error likely plays a larger role in determining whether a majority of respondents intend to vote for the election winner, than in whether they correctly forecast the winner. We will develop this insight at much greater length, in section IV.

The analysis in Table 1 does not permit strong conclusions, and indeed, it highlights two important analytic difficulties. First, we have very few national Presidential elections, and so national data will permit only noisy inferences. Second, our outcome measure—asking whether a method correctly forecasted the winner—is a very coarse measure of the forecasting ability of either approach to polling. Thus, we will proceed in two directions. First, we not turn to exploiting a much larger number of elections by analyzing data from the same surveys on who respondents expect to win the Electoral College votes of their state. And second, in sections III and IV, we will proceed to analyzing the accuracy of each approach in forecasting the two-party preferred vote share.

Forecasting the winner in each state

We begin with the state-by-state analysis, analyzing responses to the state-specific voter expectation question:

Voter Expectation (state level): *How about here in [state]. Which candidate for President do you think will carry this state?*

We compare responses to this question with the voter intention question described above. Before presenting the data, there are four limitations of these data worth noting. First, the ANES does not survey people in every state, and so in each wave, around 35 states are represented. Second, this question was not asked in the 1956-68 and 2000 election waves. We do not expect either of these issues to bias our results toward favoring either intention- or expectation-based forecasts. Third, the sample sizes in each state can be small. Across each of these state elections, the average sample size is only 38 respondents, and the sample size in a state ranges from 1 to 246. In section IV we will see that this is an important issue, as the expectation-based forecasts are relatively stronger in small samples. Fourth, while the ANES employs an appropriate sampling frame for collecting nationally representative data, it is not the frame that one would design were one interested in estimating state-specific aggregates, as these samples typically involve no more than a few Primary Sampling Units. Despite these limitations, this data still presents an interesting laboratory for testing the relative efficacy of intention versus expectation-based polling. All told we have valid ANES data from 13,208 respondents drawn from 10 election cycles (1952, 1972-1996, and 2004-2008), and in each cycle, we have data from between 28 and 40 states, for a final sample of 345 races.[1] The basic unit of observation for our forecast comparisons will be each of these 345 races, although we allow for the possibility that forecast errors may be correlated across states within an election cycle by reporting standard errors clustered by year.

Table 2 summarizes the performance of our two questions at forecasting the winning Presidential candidate in each state. Again, we use a very coarse performance metric, simply scoring the proportion of races in which the candidate who won a majority in the relevant poll ultimately won in that state—according to either our analysis of the voter forecast expectation (the first column of data), or the more standard voter intent question (the second column). When a poll yields a fifty-fifty split, we score it as half a correct call, and half an incorrect call. All told, the voter expectation question predicted the winner in 279 of these 345 races, or 80.9%, compared with 239 correct calls, or 69.3% for the voter intention question. A simple difference in proportions test reveals that these differences are clearly statistically significant ($z=3.52^{***;}$ $p<.01$). The voter expectation question outperformed the voter intent question in 8 of the 10 election cycles and tied in 2008. The difference in that tenth cycle (1972) was that Nixon won a tight race in Minnesota.

[1] Only the 311 pre-2008 elections are used in Section IV, so that we can hold the 2008 data back for true out-of-sample testing in section V.

Table 2: Forecasting the Presidential Election, by State

Year	Proportion of states where the winning candidate was correctly predicted by a majority of respondents to:		Number of states surveyed
	Expectation question	Intention question	
1952	74.3%	58.6%	35
1972	97.4%	100%	38
1976	80.3%	77.6%	38
1980	57.7%	41.0%	39
1984	86.7%	68.3%	30
1988	88.3%	56.7%	30
1992	89.4%	77.3%	33
1996	75.0%	67.5%	40
2004	89.3%	67.9%	28
2008	76.5%	76.5%	34
Totals:	**279 of 345 correct**	**239 of 345 correct**	**Difference:**
Average:	**80.9%**	**69.3%**	**11.6%*****
(Standard error)	**(3.8)**	**(5.4)**	**(3.2)**

Notes: ***, **, and * denote statistically significant coefficients at the 1%, 5%, and 10%, respectively. (Standard errors in parentheses are clustered by year).

All told, Table 2 shows the improvement in the forecasting performance obtained by relying on voter expectations is not only statistically significant, it's also quite large. The contrast is arguably even starker still, as in both questions yielded similar forecasts in 78% of these races. Thus, the real testing ground is in the remaining 22% of the sample—the 76 elections where the two questions pointed to different candidates. Of these, candidate preferred by the expectations poll won 76.3%, compared with only 23.7% of races won by the candidate who was ahead in the corresponding poll of voters intentions (and this difference is clearly statistically significant).

Another comparison comes from drawing on post-election surveys which asked who each respondent actually voted for. These ex post data (not shown) made the correct call in 70.4% of these races, which is both economically and statistically significantly worse than the proportion recorded by the voter expectation question ($z=3.16^{***}$, $p<.01$).

Thus far our analysis so far has been quite crude, in that it has evaluated only whether the favored candidate won. This approach has the virtue of transparency, but it leaves much of the variation in the data—such as variation in the winning margin—unexamined. Thus we now turn to analyzing the accuracy of forecasts of vote shares derived from both intention and expectation data. We will also add some structure to how we are thinking about these data. At this point we drop 2008 from the dataset, in order to have some "out of-sample" data to review in Section V.

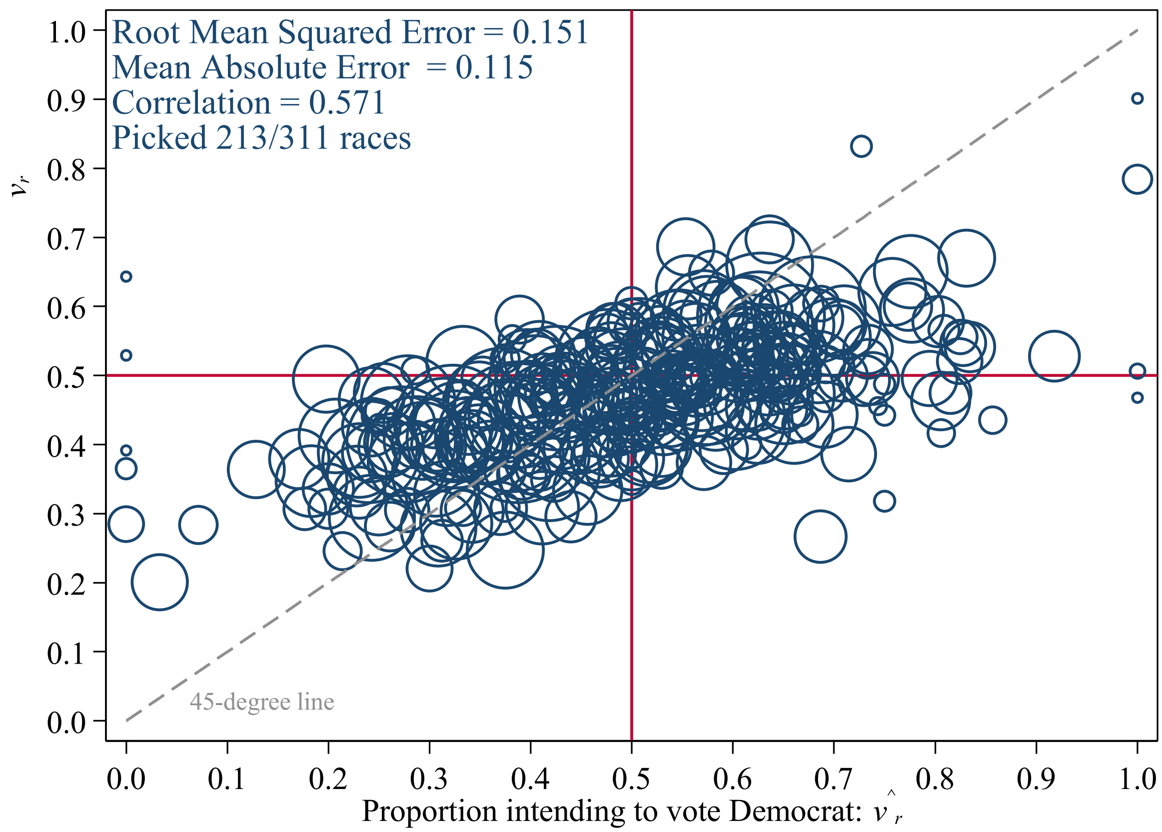
III. Simple (or Naïve) Forecasts of the Vote Share

Our goal is to use the state-by-state ANES data to come up with forecasts of the two-party vote share in each of the state×year races in our dataset. To introduce our notation, we denote **v**oting intentions as v, and voter's e**x**pectations as x, and will use the subscript r to denote a specific state-year **r**ace.

In this section, we begin by analyzing our data the way polling numbers are typically used, interpreting a survey that says that $\widehat{v_r}$ percent of *sample* respondents will vote for a candidate in state-year race r as a forecast that this candidate will win v_r percent of votes among the entire *population*. That is, we follow the norm among pollsters and make our projections as if the sample moments represent population proportions. Likewise, we interpret a poll that says that $\widehat{x_r}$ percent of sample respondents expect a candidate to win as a forecast that x_r percent of the population expect that candidate to win. In fact, raw polling data rarely represent optimal forecasts. Thus, we refer to the projections in this section as "naïve forecasts", and in section IV we will derive and evaluate statistically efficient forecasts.

Our focus is on predicting each candidate's share of the two-party vote, and we begin by analyzing data on voters' intentions. Figure 1 plots the relationship between the actual Democratic vote share in each state-year race, and the proportion of poll respondents who plan to vote for the Democratic candidate. There are two features of these data to notice. First, election outcomes and voter expectations are clearly positively correlated—that is, these polls are informative. But second, the relationship is by no means one-for-one, and these relatively small polls of voter intention are only a noisy measure of the true vote share on Election Day.

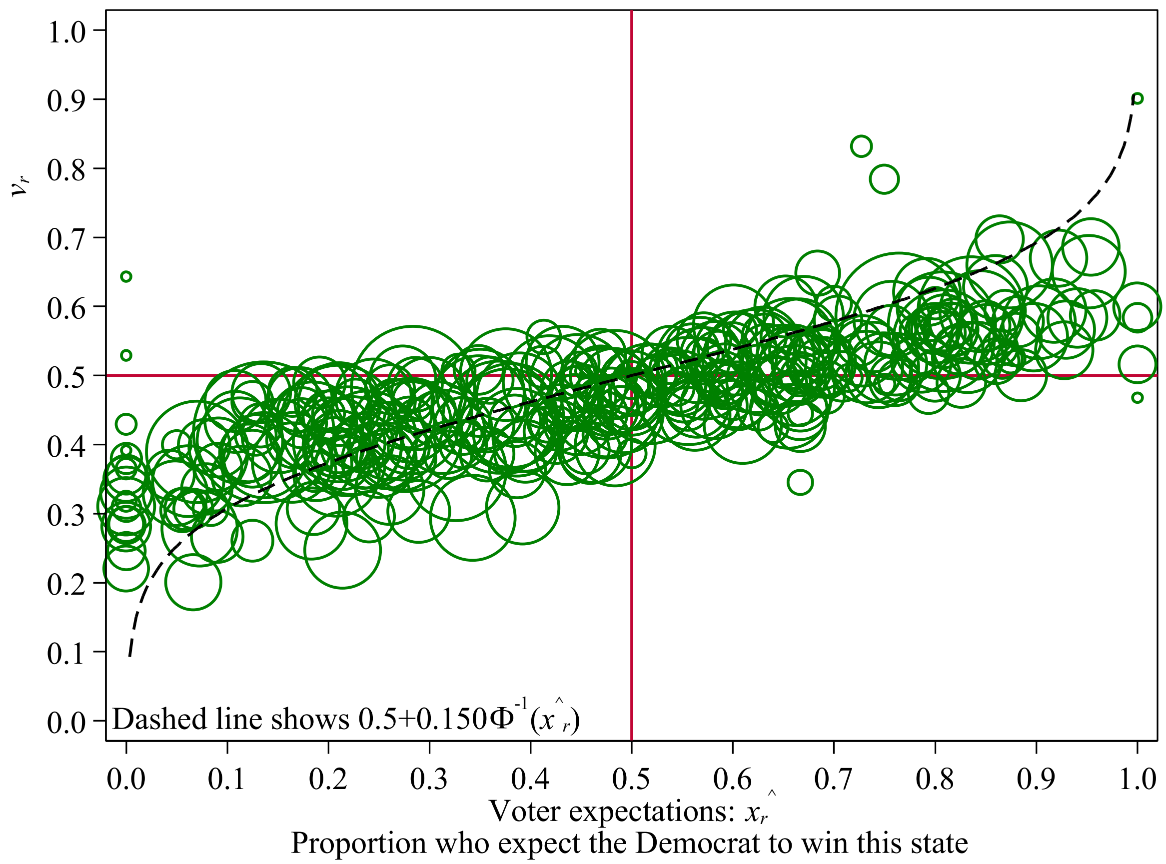
Figure 1: Naïve Voter Intention Forecast and Actual Vote Share



Notes: Each point shows a separate state-year cell in a Presidential Electoral College election; the size of each point is proportional to the number of survey respondents. Both voter intention and election outcomes refer to shares of the total votes cast for the two major parties. There are a total of n=311 elections, as the 2008 data is not included.

In Figure 2 we show the relationship between voter expectations—the share of voters who expect the Democrat to win that state's presidential ballot—and the vote share he actually garnered. This plot reveals that there is a close relationship between election outcomes and voter expectations, and typically the candidate who most respondents expect to win, does in fact win. That is, most of the data lie in either the northeast or southwest quadrants, a fact also evident in Table 2. Equally, the relationship between voter expectations and vote shares does not appear to be linear. This isn't surprising, as a statement that two-thirds of voters expect Obama to win is not a statement that he'll win two-thirds of the vote. As such, we will now add some structure with an eye to teasing out the forecast of vote shares implicit in these data.

Figure 2: Voter Expectation and Actual Vote Share



Notes: Each point shows a separate state-year cell in a Presidential Electoral College election; the size of each point is proportional to the number of survey respondents. The proportion who expect the Democrat to win and the corresponding vote share refer to shares of the expectations and total votes cast for the two major parties, respectively. The dashed line shows the mapping from the proportion of the population who expect a candidate to win to the corresponding vote share implied by equation [6]. There are a total of n=311 elections, as the 2008 data is not included.

We begin by characterizing how people respond to the question probing their expectations about the likely winner. If each person views an unbiased noisy signal, ${x^*}_r^i$ of a candidate's final vote share—where the superscript i serves as a reminder that we are analyzing individual responses, and the asterisk reminds us that this is an unobserved latent variable—then:

$$ {x^*}_r^i = v_r + \epsilon_r^i \text{ and } \epsilon_r^i \sim N(0, \sigma_\epsilon^2) \qquad [1] $$

where ϵ_r^i is an idiosyncratic error reflecting the respondent's imperfect observation.[2] We assume that this noise term is drawn from a normal distribution, and its variance is constant across both poll respondents, and across elections (and to the extent this over-simplifies, it will lead us to understate the potential accuracy of expectations-based forecasts). In turn, if people expect a specific candidate to win if this noisy signal says that this candidate will win at least half the vote, then they will respond to questions about voter expectations as follows:

$$x_r^i = \begin{cases} 1 \quad if\ {x^*}_r^i = v_r + \epsilon_r^i > 0.5 \\ 0 \quad if\ {x^*}_r^i = v_r + \epsilon_r^i < 0.5 \end{cases} \qquad [2]$$

Consequently equations [1] and [2] together imply that we can estimate σ_ϵ from a simple probit regression explaining whether the respondent expected a candidate to win, by a variable describing the extent by which the vote share garnered by that candidate exceeds the 50% required to win:

$$E[x_r^i = 1|v_r] = \Phi\left(\frac{1}{\sigma_\epsilon}(v_r - 0.5)\right) \qquad [3]$$

This regression yields an estimate of $1/\widehat{\sigma_\epsilon} = 6.661$ with a standard error allowing for within-state-year correlated errors of 0.385 (n=11,548), which implies that $\widehat{\sigma_\epsilon} = 0.150,$ with a standard error of 0.0089 (estimated using the delta method). This estimate suggests that the typical respondent is quite uninformed about the likely result of the upcoming election. For now, we simply note that this estimate provides a link between election results, and the proportion of the population who expect the Democrat to win, x_r:

$$E[x_r|v_r] = Prob\left(v_r + \epsilon_r^i > 0.5\right) = \Phi\left(\frac{v_r - 0.5}{\sigma_\epsilon}\right) \qquad [4]$$

When the voting population is large then the noise induced by ϵ^i in the mapping between the population parameters v_r and x_r is negligible,[3] and so it follows that:

$$x_r \approx \Phi\left(\frac{v_r - 0.5}{\sigma_\epsilon}\right) \qquad [5]$$

From this, we can back out the implied expected vote share by inverting this function:

[2] Part of this error may be due to the fact that the election is still a month away; thus ϵ^i includes variation due to voters who may later change their minds.

[3] As we will see in the next section, the noise term ϵ_r^i is relevant to the mapping between the population parameter x_r and the sample estimate $\widehat{x_r}$.

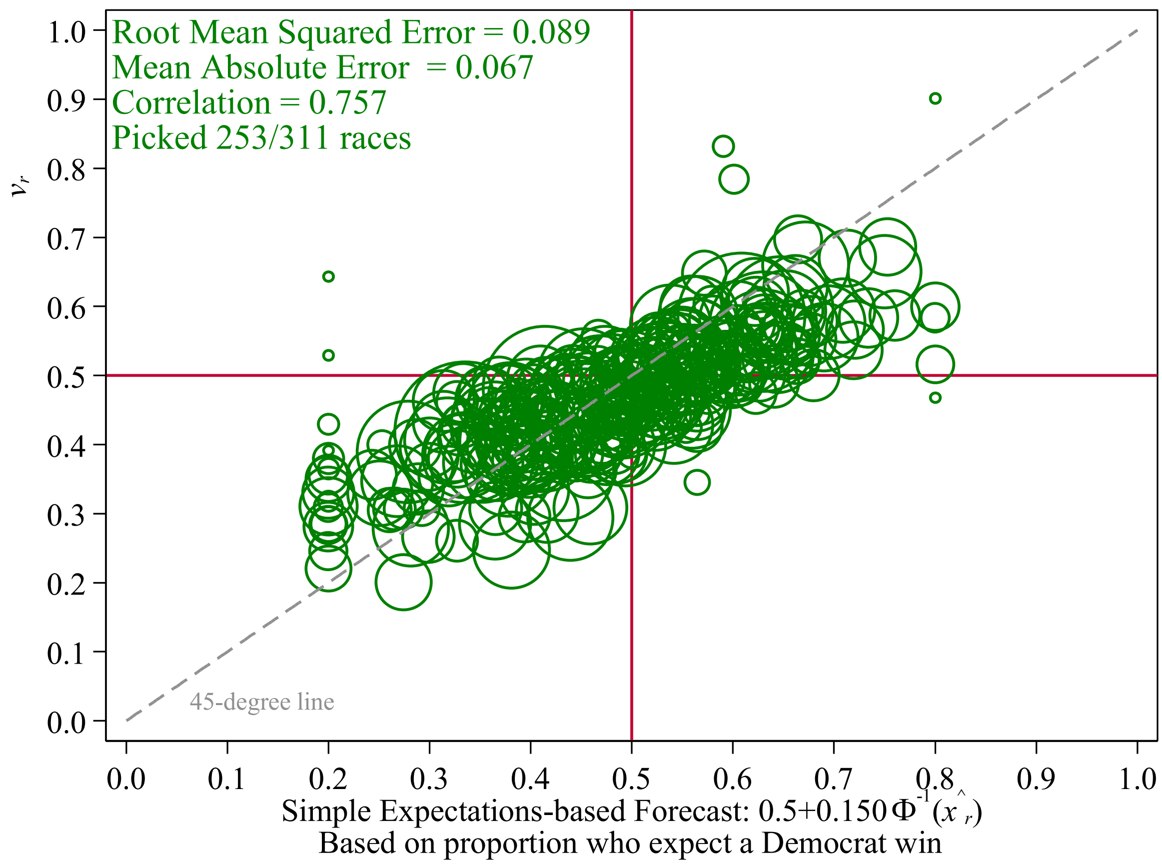
$$E[v_r|x_r] \approx 0.5 + \sigma_\epsilon \Phi^{-1}(x_r) \qquad [6]$$

This function is also shown as a dashed line in Figure 2, based on our estimated value of $\widehat{\sigma_\epsilon} = 0.150$. To be clear, this is the appropriate mapping only if we know the true population proportion who expect a particular candidate to win, x_r. While this assumption is clearly false, our goal here is to provide a forecast comparable to the naïve forecast of voter intentions, and so in both cases we use the mapping between survey proportions and forecasts that would be appropriate in the absence of sampling variation. Indeed, in Figure 2 many of the extreme values of the proportion of voters expecting a candidate to win likely reflect sampling variation. That is, our transformation of voter expectations data into vote shares is *not* estimated as a line of best fit between aggregate vote shares and expectations. Indeed, elections are rarely as lopsided as the dashed line suggests. This feature roughly parallels the observation that in Figure 1; elections are rarely as lopsided as suggested by small samples of voter intentions. We will explore this feature of the data in greater detail in the next section when we evaluate efficient shrinkage-based estimators. But for now we will call this simple transformation of voter expectation our "naïve" expectation-based forecast.[4]

Figure 3 plots our naïve expectations-based forecasts of vote shares against the actual election results. These forecasts are clustered along the 45-degree line, suggesting that they are quite accurate.

[4] The one remaining difficulty is that in 22 races (7% of them), either 0% or 100% of survey respondents expect the Democrat to win, and so equation [6] does not yield a specific forecast. In these cases, we infer that the candidate is expected to win 20% or 80% of the vote, respectively. Our admittedly arbitrary rational is simply that these are the nearest round numbers that ensure the implied forecast is monotonic in the proportion of respondents expecting a particularly candidate to win. We obtain qualitatively similar results when either imputing expected vote shares of 0% and 100% instead, or simply dropping these cases from the sample. Section IV provides a more satisfactory treatment of this issue.

Figure 3: Naïve Expectation-Based Forecast and Actual Vote Share



Notes: Each point shows a separate state-year cell in a Presidential Electoral College election; the size of each point is proportional to the number of survey respondents. Both the expectations-based forecast and the actual vote outcomes refer to shares of the total votes cast for the two major parties. There are a total of n=311 elections, as the 2008 data is not included.

In Table 3 we provide several simple comparisons of forecast accuracy. The first two rows show that the expectation-based forecast yields both a root mean squared error and mean absolute error that is significantly less than the intention-based forecast. The third row shows that the expectation-based forecast is also the more accurate forecast in 65% of these elections. The final column shows that in each case these differences are clearly statistically significant. In the fourth row, we examine the correlation coefficient. One might be concerned that the better performance of the expectation-based forecasts reflects the fact that they rely on an estimated parameter, σ_ϵ, and thus they use up one more degree of freedom. That is, our estimated value of σ_ϵ "tilts" the expectations data so that the implied forecasts lie along the 45-degree line. The correlation coefficient effectively both tilts the data and shifts it up and down, so as to maximize fit. Thus, it arguably puts each forecast on something closer to an equal footing.

Even so, the expectation-based forecasts are also more highly correlated with actual vote shares than are the intention-based forecasts. [5]

Table 3: Comparing the Accuracy of Naïve Forecasts of Vote Shares

	Raw Voter Intention: $\widehat{v_r}$	Transformed Voter Expectation: $0.5+0.150\Phi^{-1}(\widehat{x_r})$	Test of Equality
Root Mean Squared Error	0.151 (0.014)	0.089 (0.007)	t_8=4.96 (p=0.001)
Mean Absolute Error	0.115 (0.009)	0.067 (0.004)	t_8=8.49 (p<0.001)
How often is forecast closer?	35.0% (1.6)	65.0% (1.6)	t_8=9.43 (p<0.001)
Correlation	0.571	0.757	
Encompassing regression: $v_r = \alpha + \beta_v Intention_r + \beta_x Expectation_r$	0.058 (0.055)	0.480*** (0.023)	$F_{1,8}$=39 (p<0.001)
Optimal weights: $v_r = \beta Intention_r + (1-\beta) Expectation_r$	8.5%* (4.5)	91.5%*** (4.5)	$F_{1,310}$=84.1 (p<0.001)

Notes: ***, **, and * denote statistically significant coefficients at the 1%, 5%, and 10%, respectively. (Standard errors in parentheses, clustered by year). These are assessments of forecasts of the Democrat's share of the two-party vote in n=311 elections. Comparisons in the third column test the equality of the measures in the first two columns. In the encompassing regression, the constant $\hat{\alpha} = 0.207^{***}$ (se=0.028).

We also show a Fair-Shiller (1989 and 1990) regression, attempting to predict election outcomes on the basis of a constant, and our two alternative forecasts. The expectation-based forecast has a large and extremely statistically significant weight, as does the constant. Moreover it fully encompasses the information in the intention-based forecast, in that conditional on our expectations-based forecast, the intentions-based forecast receives a statistically insignificant (and small) weight. Finally, we estimate the optimal weighted average of these two forecasts, which puts greater than 90% of the weight on the expectation-based forecast.

Our assessment so far is entirely based on the common but problematic assumption that our sample data of voter intentions and expectations can be interpreted as representing population moments. We now turn to generating statistically efficient forecasts instead, and will repeat our forecast evaluation exercise based on these adjusted forecasts.

[5] The findings in Table 3 are remarkably similar when we produce our expectations-based forecast for each race using an estimate of σ_ϵ derived from a sample which excludes that race.

IV. Efficient Forecasts of the Vote Share

An important insight common to both Figure 1 and Figure 3 is that in those elections where the Democrats are favored, the final outcome typically does favor Democrats, but by less than suggested by the naïve forecasts based on either voter intentions or voter expectations. Likewise, when the poll favors Republicans, the Democrats do tend to lose, but again, by less than suggested by our naïve forecasts. That is, election outcomes are typically closer than suggested by either of our naïve forecasts. In fact, this is a natural implication of sampling error: the more extreme polling outcomes likely reflect sampling variation, and so it is unsurprising that they are not matched by equally extreme electoral outcomes. It is this observation that motivates our use of shrinkage estimators in this section, "shrinking" the raw estimates of the proportion intending to vote for one candidate or another toward a closer race. This idea is widely understood by political scientists (Campbell 2000), but is typically ignored in media commentary. We will also adjust for any biases (or "house effects") in these data.

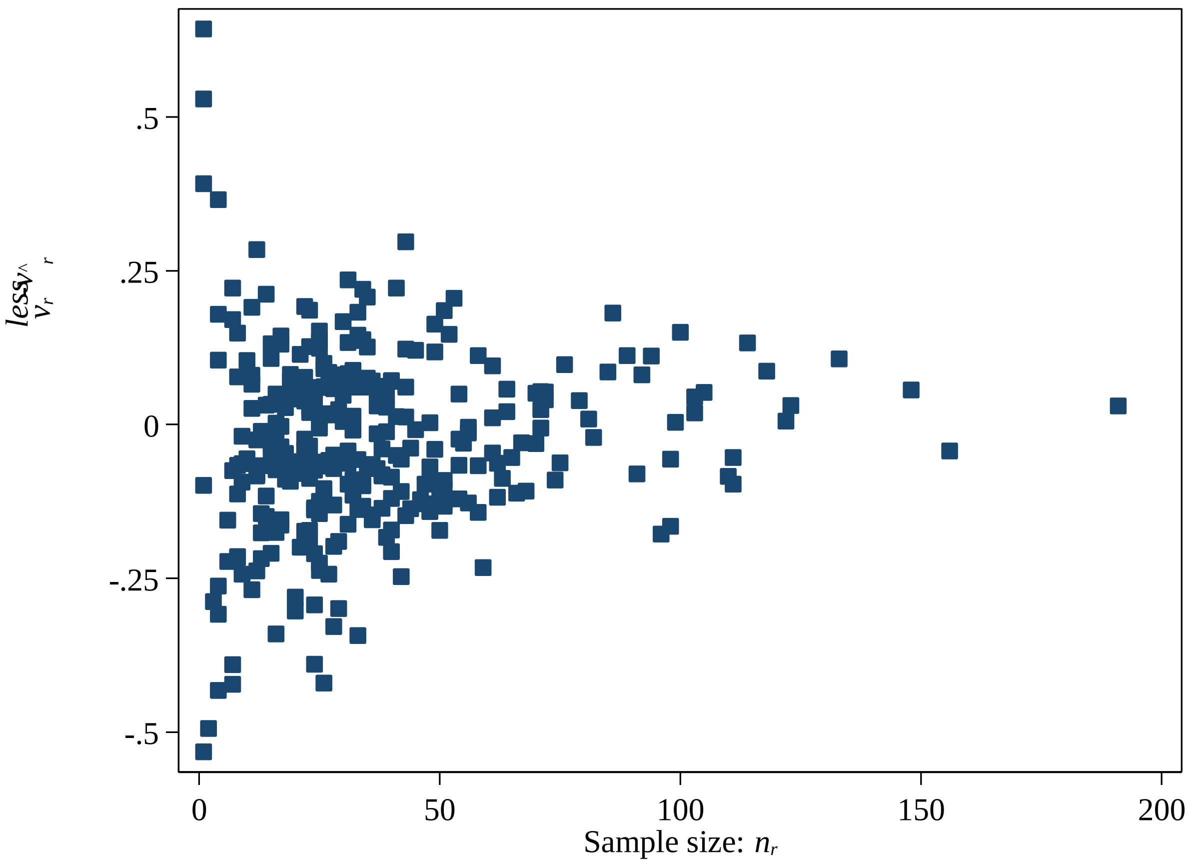
In the following discussion it is important to distinguish between the actual vote share won by the Democrat, v_r, from the sample proportion who intend to vote for the candidate, $\widehat{v_r}$, and the optimal intention-based forecast based on polls of voter intentions, $E[v_r|\widehat{v_r}]$. Likewise, we distinguish the sample proportion who expect a candidate to win, $\widehat{x_r}$, from the population proportion, x_r, and our optimal expectation-based forecast of the vote share, $E[v_r|\widehat{x_r}]$. Because we are only analyzing respondents with valid expectation and intent data, each forecast will be based on the same sample size, n_r.

We will begin by analyzing forecasts based on standard polls of voter intentions, and will then turn to analyzing how voter expectations might improve these forecasts.

Interpreting Voter Intentions

Our goal is to find the mapping between our raw voter intentions data, and the forecast which minimizes the mean squared forecast error. The intuition of our approach can be gleaned from thinking about an OLS regression predicting the final vote share based on polling results. The problem is that polling results are subject to sampling error. When there are errors in the explanatory variable, the least squares coefficient will be shrunk by a factor related to the signal-to-noise ratio in that variable, and the constant plays a bigger role. While that intuition applies here, our case is complicated by heteroscedasticity. That is, the sample sizes of our polls vary widely across each race, and as Figure 4 illustrates, so too does the noise underpinning each observation. In turn this implies that the extent to which each poll reflects signal versus noise also varies. Consequently the extent to which the sample estimate should be shrunk towards the grand mean differs for each race, and it does so in a way related to the sample size.

Figure 4: Sample Size and Forecast Errors in the Intent Poll



Let's now formalize this intuition. The key issue that we address in this section is that our sample estimate of the proportion of voters intending to vote for a candidate is a noisy and possibly also biased estimate of the election outcome:

$$\widehat{v_r} = v_r + \alpha^v + (\eta_r + \tau_r), \text{ where } \eta_r \sim N(0, \sigma^2_{\eta_r}), \tau_r \sim N(0, \sigma^2_\tau) \text{ and } E[\eta_r \tau_r] = 0 \qquad [7]$$

where α_v is a bias term which picks up the pro-Democrat house effect in ANES polls of voter intentions.[6] There are also two (uncorrelated) sources of polling error. First, η_r reflects sampling variability; notice the r subscript on the variance $\sigma^2_{\eta_r}$, which reflects the fact that sampling variability will vary with the sample size of each specific poll. Second, τ_r is an uncorrelated error term, reflecting the fact that we are sampling from a population where v_r percent of respondents will ultimately vote for the Democrat, but

[6] We also tested for an anti-incumbent party bias, but found it to be small and statistically insignificant.

when they are polled one month prior, an extra τ_r percent intend to vote Democratic. Assuming that the total polling error is orthogonal to the actual election result ($E[v_r(\tau_r + \eta_r)] = 0$), we get the following familiar result:

$$E[v_r|\widehat{v_r}] = \mu^v + \frac{\sigma_v^2}{\sigma_v^2 + \sigma_{v_r - \widehat{v_r}}^2} (\widehat{v_r} - \alpha^v - \mu^v) \quad [8]$$

where μ^v and σ_v^2 are the mean and variance of the Democratic vote share, across all the races in our dataset. In order to use this equation to formulate our efficient intentions-based forecast, we need to estimate each of the greek-letter parameters. We estimate the average vote share of Democrats directly from our sample of R election races: $\widehat{\mu^v} = \frac{1}{R}\sum v_r = 0.468$ $(se = 0.005)$, and the variance of the Democrat vote share is: $\widehat{\sigma_v^2} = \frac{1}{R-1}\sum(v_r - \widehat{\mu_v})^2 = 0.0089$. Likewise, it is easy to estimate the bias term, $\widehat{\alpha^v} = \sum(\widehat{v_r} - v_r)/R = 0.031$ $(se = 0.008)$. This bias in the ANES is reasonably large, statistically significant, and to our reading, has not previously been documented.[7] All that remains is to sort out the variance of the polling errors, $\sigma_{v_r - \widehat{v_r}}^2$, which can be expressed:

$$\sigma_{v_r - \widehat{v_r}}^2 = \sigma_\tau^2 + \sigma_{\eta_r}^2 \quad [9]$$

The first term reflects the variability in "true" public opinion between polling day and Election Day. For a subset of our sample, we observe both their voting intentions, and how they actually voted, and so we can estimate this as $\widehat{\sigma_\tau^2} = 0.00035$.[8] The second term in equation [9] reflects sampling variability, which depends on the sample size in that particular poll, and hence earns an r subscript. Because the poll result $\widehat{v_r}$ is the mean of a binomial variable with mean $v_r + \alpha^v + \tau_r$, this second term can be expressed as:

$$\sigma_{\eta_r}^2 = \frac{(v_r + \alpha^v + \tau_r)\big(1 - (v_r + \alpha^v + \tau_r)\big)}{n_r^{eff}} \approx \frac{\frac{1}{4}}{n_r/\gamma_r^v} = \frac{\gamma_r^v}{4n_r} \quad [10]$$

The numerator of this expression is the product of the poll shares of the two parties, if the election were held on polling day. For most elections (and particularly competitive contests), the term $(v_r + \alpha^v + \tau_r)\big(1 - (v_r + \alpha^v + \tau_r)\big) \approx \frac{1}{4}$, and so we use this approximation. (We obtain similar results when we

---

[7] This pro-Democrat bias in the ANES remains statistically significant even when we cluster results by year.

[8] While many people change their vote between being polled and election day, in large electorates what matters is common changes in voting intentions. Thus, we turn to the micro data, and estimate the following random effects model: $I(vote\ Democrat_r^i) = I(intend\ to\ vote\ Democrat_r^i) + \tau_r + \eta_r$. We use the Swamy-Arora estimator of σ_τ^2, which includes a small-sample adjustment for unbalanced panels, and find $\widehat{\sigma_\tau^2} = 0.00035$. This estimate is comparable to Lock and Gelman, who estimate σ_τ^2 as a function of time before the election, exploiting the fact that polls become more accurate as election day approaches.

plug in actual vote shares or vote shares from the previous election instead; our approximation has the virtue of being usable in a real-time forecasting context.) The denominator, n_r^{eff} denotes the effective sample size of the specific poll in race r. If the sample were a simple random sample, the effective sample size would be exactly equal to the actual sample size. But the American National Election Studies uses a complex sample design, polling only in a limited number of primary sampling units. The "design effect" γ_r^v corrects for the effects of the intra-cluster correlation within these sampling units. (The subscript r serves as a reminder that it varies with the sample size of specific poll—for instance, it is one when $n_r = 1$—and the superscript v serves as a reminder that the design effect varies, and this is the design effect for voter intentions). Unfortunately published estimates of γ_r^v are based on design effects in national samples, and so they can't be applied to our analysis of state samples. Moreover, the public release files in the ANES files do not contain sufficient detail about the sampling scheme to allow us to estimate these design effects directly. It's easy to show that the design effect here is $\gamma_r^v = 1 + \frac{\rho^v}{1-\rho^v} n_r$, where ρ^v is the intra-cluster correlation coefficient for voter intentions. In what follows, we assume that ρ^v is constant across states and time.[9] While we lack the details on sampling clusters to estimate ρ^v directly, we can estimate it indirectly. Figure 4 highlights the underlying variation that identifies ρ^v, illustrating how polling errors varying with the sample size of the poll. Our identification comes from the fact that this pattern is shaped by ρ^v—the higher is ρ^v, the less quickly the variance of $(v_r - \widehat{v_r})$ declines with sample size. Thus, we return to equation [8], plug in the values for $\widehat{\mu^v}$, $\widehat{\sigma_v^2}$, $\widehat{\sigma_\tau^2}$ and $\widehat{\alpha^v}$ and estimate ρ^v directly by running non-linear least-squares on the following regression:

$$v_r = \widehat{\mu^v} + \frac{\widehat{\sigma_v^2}}{\widehat{\sigma_v^2} + \widehat{\sigma_\tau^2} + \frac{1 + n_r \rho^v/(1-\rho^v)}{4 n_r}} \left(\widehat{v_r} - \widehat{\alpha^v} - \widehat{\mu^v}\right) \qquad [11]$$

which yields an estimate of $\widehat{\rho^v} = 0.031$ (with a standard error of 0.0074), implying an average design effect of $\widehat{\gamma^v} = 2.18$. Thus, returning to equation [8], our MSE-minimizing forecast based on the voter intentions data, is:

$$E[v_r | \widehat{v_r}] = \mu^v + \left[\frac{\sigma_v^2}{\sigma_v^2 + \sigma_{v-\hat{v}}^2}\right] \left(\widehat{v_r} - \alpha^v - \mu^v\right) \qquad [12]$$

[9] Our assumption that this correlation is constant across states is surely problematic, given that our estimates of the intra-cluster correlation are estimates across the set of sampled addresses within a state (not estimates within a sampling unit) and most small state samples reflect only a single cluster, while some of the larger states include multiple primary sampling units. But we persist with this assumption because our data does not consistently code sufficient data on the sampling scheme used.

$$= 0.468 + \left[\frac{0.0089}{0.0089 + 0.00035 + \frac{1 + n_r 0.031/(1 - .031)}{4n_r}} \right] (\hat{v_r} - 0.031 - 0.468)$$

Given the data in our sample, the shrinkage term [in square brackets] averages 0.33 (which corresponds closely with the average slope seen in Figure 1, but it ranges from 0.03 (in a race with only one survey respondent) to 0.48 (in a race with 191 respondents), and with an infinitely large sample, approaches 0.52.

Figure 5: Efficient Intention-Based Forecasts and Actual Vote Share

In Figure 5, we show the relationship between our optimal intention-based forecast and actual vote share. These adjusted intention-based forecasts are clearly more accurate than the naïve forecasts numbers: the forecasts lie along the 45-degree line, and both the mean absolute error and the root mean squared error are about half that found Figure 1. We now turn to finding the most efficient forecasts of vote shares based on voter expectations.

Interpreting Voter Expectations

In our previous analysis in Section III, we transformed data on voter expectations into vote share forecasts based on $E[v_r|x_r]$, the relationship between vote shares and expectations in the whole population. But taking the sample variability seriously means that we need to find the appropriate mapping, based on a sample of expectations data, $E[v_r|\widehat{x_r}]$. Using equation [6], we can express this as $E[0.5+\sigma_\epsilon\Phi^{-1}(x_r)|\widehat{x_r}]$, which is a non-linear function with no closed-form solution.[10] However note that for elections where there is substantial disagreement about the likely outcome—which ensures that $\Phi^{-1}(x_r)f(x_r|\widehat{x_r})$ is nearly linear—we can use the approximation:

$$E[v_r|\widehat{x_r}] = E[0.5+\sigma_\epsilon\Phi^{-1}(x_r)|\widehat{x_r}] \approx 0.5+\sigma_\epsilon\Phi^{-1}(E[x_r|\widehat{x_r}]) \quad [13]$$

The key input to equation [13] is $E[x_r|\widehat{x_r}]$, and so once again, we turn to a shrinkage estimator in order to generate efficient estimates of x_r, given our small sample estimates, $\widehat{x_r}$. Following the same logic we used (in equation [8]) for generating the efficient estimator of voter intentions, the efficient shrinkage estimator is:[11]

$$E[x_r|\widehat{x_r}] = \mu^x + \left[\frac{\sigma_x^2}{\sigma_x^2+\sigma_{x_r-\widehat{x_r}}^2}\right](\widehat{x_r}-\alpha^x-\mu^x) \quad [14]$$

where μ^x is the mean across all elections of the proportion of the population who expect the Democrat to win; σ_x^2 is the corresponding variance, while $\sigma_{x_r-\widehat{x_r}}^2$ is the variance of the corresponding sampling error; and as before, we have a bias parameter, α^x which allows for the possibility that these data oversample people who expect Democrats to win (although it isn't directly comparable to the bias in voter intentions, α^v).

The key difficulty in working with data on voter expectations rather than voter intentions is that while we do ultimately observe how the entire population votes, we never observe what the whole population expects. Thus in estimating population parameters, we will rely heavily on the mapping in

---

[10] The exact solution is $E[v_r|\widehat{x_r}] = 0.5+\sigma_\epsilon\int_0^1\Phi^{-1}(x)f(x|\widehat{x_r})dx$. In turn, this depends on $f(x|\widehat{x_r})$, the probability density function of the proportion of the population who expect the Democrat to win, given the observed sample expectations. This is recovered from Bayes' Rule: $f(x|\widehat{x_r}) = \frac{L(\hat{x}_r|x)p(x)}{\int_0^1 L(\widehat{x_r}|x\prime)p(x\prime)dx\prime}$ where the likelihood, $\mathrm{L}(\widehat{x_r}|x)\sim Binomial(n_r,\widehat{x_r}n_r,x)$ and the prior, $p(x)$ comes from the earlier observations that $v_r\sim N(0.468,.0089)$ and $x=\Phi\left(\frac{v_r-0.5}{\sigma_\epsilon}\right)$. The correlation between the approximation we use in the text and this numeric solution is 0.9980.

[11] Technically, this is appropriate only if x is normally distributed, which would be unlikely given the distributional assumptions made in equations [1]-[3]. But see the previous footnote for the demonstration that this approximation seems reasonable, in practice.

equation [6] between population vote shares and population expectations. Using this insight, we estimate $\widehat{\mu^x} = \sum \Phi\left(\frac{v_r - 0.5}{\widehat{\sigma_\epsilon}}\right)/R = 0.427\ (se = 0.011)$. Likewise we estimate the bias term $\widehat{\alpha^x} = \sum (\widehat{x_r} - \Phi\left(\frac{v_r - 0.5}{\widehat{\sigma_\epsilon}}\right))/R = 0.043\ (se = 0.009)$, which represents the higher probability that a survey respondent expects the Democrat to win, relative to the population mean.

The numerator of the shrinkage estimator in equation [14] is the variance of the population expectation across all elections, and it is also quite straightforward to estimate: $\widehat{\sigma_x^2} = \frac{1}{R-1}\sum\left(x_r - \widehat{\mu^x}\right)^2 = 0.040$. All that remains is to sort out the sampling variability of our intentions data, $\sigma^2_{x_r - \widehat{x_r}}$. Because different elections have different sample sizes, this term varies across elections. Again, because we are asking about binary outcome—whether or not you expect the Democrat to win in your state—we know the functional form of the relevant sampling error. Thus:

$$\sigma^2_{x_r - \widehat{x_r}} = \frac{x_r(1-x_r)}{n_r^{eff}} \approx \frac{\gamma_r^x}{4n_r} \qquad [15]$$

where the approximation follows because the product of the population proportions expecting each candidate, $x_r(1-x_r) \approx \Phi\left(\frac{v_r - 0.5}{\widehat{\sigma_\epsilon}}\right)\Phi\left(\frac{0.5 - v_r}{\widehat{\sigma_\epsilon}}\right)$, and in turn, $\Phi\left(\frac{v_r - 0.5}{\widehat{\sigma_\epsilon}}\right)\Phi\left(\frac{0.5 - v_r}{\widehat{\sigma_\epsilon}}\right) \approx \frac{1}{4}$ because most elections are competitive (and $\widehat{\sigma_\epsilon}$ is not too small).[12] As before, n_r^{eff} is the effective sample size, and the relevant design effect is $\gamma_r^x = 1 + \frac{\rho^x}{1-\rho^x} n_r$.

Putting equations [13], [14] and [15] together, yields:

$$E[v_r|\widehat{x_r}] \approx 0.5 + \widehat{\sigma_\epsilon}\Phi^{-1}\left(\overbrace{\widehat{\mu^x} + \frac{\widehat{\sigma_x^2}}{\widehat{\sigma_x^2} + \frac{1 + n_r\rho^x/(1-\rho^x)}{4n_r}}\left(\widehat{x_r} - \widehat{\alpha^x} - \widehat{\mu^x}\right)}^{E[x_r|\widehat{x_r}]}\right) \qquad [16]$$

At this point, the only parameter in this equation that remains unknown is the intra-correlation coefficient ρ^x (where the superscript reminds us that there is no reason to expect the intra-cluster correlation in voter expectations to be similar to that for voter intentions.) As we did with the voter intentions data, we use an indirect approach, plugging our estimates of the parameters $\widehat{\sigma_\epsilon}, \widehat{\mu_x}, \widehat{\sigma_x^2}, \widehat{\alpha^x}$ and

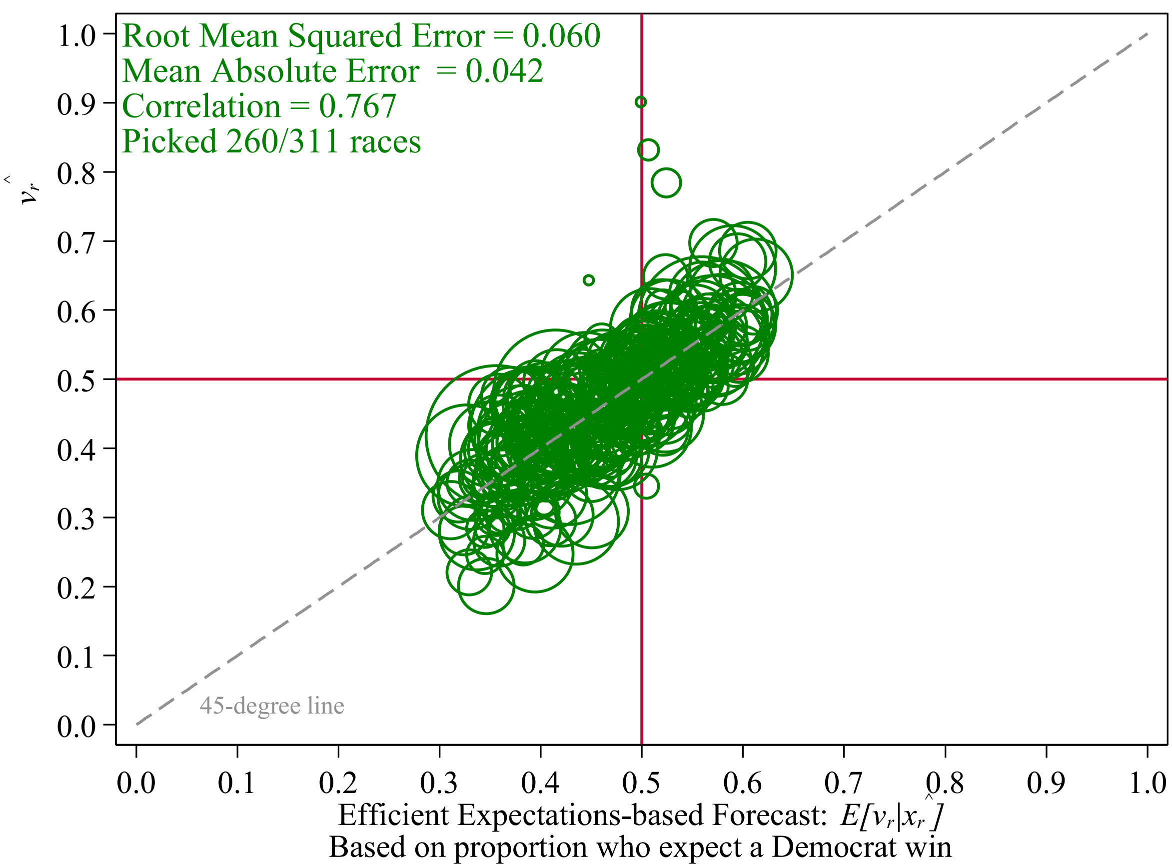
[12] In our small samples, $\widehat{x_r}(1 - \widehat{x_r})$ can diverge quite significantly from 0.25 but equation [14] involves the product of the *population* proportions who expect each candidate to win. While we don't actually observe these population proportions, but we do observe the population vote shares, and using the transformation in equation [6], have confirmed that $\Phi\left(\frac{v_r - 0.5}{\widehat{\sigma_\epsilon}}\right)\left(1 - \Phi\left(\frac{0.5 - v_r}{\widehat{\sigma_\epsilon}}\right)\right) \approx 0.25$.

the data $\widehat{x_r}$ and n_r back into equation [16], which allows us to estimate ρ^x using a non-linear least squares regression. This yields an estimate of $\widehat{\rho^x} = 0.041$ (se $=$ 0.009), which in turn implies an average design effect $\widehat{\gamma^x} = 2.60$, and that the shrinkage term which has an average value of 0.65, ranges from 0.13 to 0.77. Thus, our MSE-minimizing forecast of the Democrats vote share, based only the voter expectations data is:

$$E[v_r|\widehat{x_r}] = 0.5 + 0.150\Phi^{-1}\left(0.427 + \frac{0.040}{0.040+\frac{1+n_r\times 0.041/(1-0.041)}{4n_r}}(\widehat{x_r} - 0.043 - 0.427)\right) \qquad [17]$$

In Figure 6, we show the relationship between our efficient expectation-based forecast and actual election outcomes. These adjusted expectation-based forecasts are clearly very accurate, and appropriately scaled: the forecasts lie along the 45-degree line.

Figure 6: Efficient Expectation-Based Forecast and Election Outcomes



Forecast evaluation

In Table 4, we present formal tests of the relative accuracy of the efficient forecasts derived from our data on voter intentions and expectations. These tests demonstrate that our forecasts based on efficiently exploiting voter expectations data are more accurate than those based on voter intentions data. The first two rows show that the expectation-based forecasts yield both a root mean squared error and mean absolute error that is smaller than the intention-based forecasts by a statically significant amount. The third row shows that the expectation-based forecasts are also the more accurate forecast in 63% of these elections, very similar to our evaluation of the simpler (naïve) forecasts. The efficient expectation-based forecasts are also more highly correlated with actual vote shares than are the corresponding intention-based forecasts, and the difference is both significant and sizable margin. In the Fair-Shiller regression, the expectation-based forecasts is statistically significant, while the intentions-based forecast is not, allowing us to infer that the former encompasses the latter, which means that we failed to find any evidence that there is any information in the intentions-based forecast that is not in the expectations-based forecast. Finally, an optimally weighted average still puts more than 90% of the weight on the expectation-based forecast. It is worth contrasting this with current polling practice, in which the average weight placed on expectation polls by pollsters, the press, and campaigns is effectively 0%, as expectations data are generally ignored.

Table 4: Comparing the Accuracy of Efficient Forecasts of Vote Share

	Efficient Voter Intention: $E[v_r \mid \widehat{v_r}]$	Efficient Voter Expectation: $E[v_r \mid \widehat{x_r}]$	Test of Equality
Root Mean Squared Error	0.076 (0.006)	0.060 (0.006)	t_8=6.20 (p<0.001)
Mean Absolute Error	0.056 (0.004)	0.042 (0.003)	t_8=6.54 (p<0.001)
How often is forecast closer?	37.0% (2.4)	63.0% (2.4)	t_8=5.36 (p=0.001)
Correlation	0.592	0.767	
Encompassing regression: $v_r = \alpha + \beta_v Intention_r + \beta_x Expectation_r$	0.182 (0.127)	0.912*** (0.036)	$F_{1,8}$=27.2 (p<0.001)
Optimal weights: $v_r = \beta Intention_r + (1-\beta) Expectation_r$	9.6%** (4.1)	90.4%*** (4.1)	$F_{1,310}$=96.2 (p<0.001)

Notes: ***, **, and * denote statistically significant coefficients at the 1%, 5%, and 10%, respectively. (Standard errors clustered by year in parentheses). These are assessments of forecasts of the Democrat's share of the two-party vote in n=311 elections. Comparisons in the third column test the equality of the measures in the first two columns. In the encompassing regression, the constant $\hat{\alpha} = -0.045$ (se=0.06).

V. Out-of-Sample Tests: 2008 Election

We now provide an out-of-sample comparison, testing how the framework developed above and the parameters estimated using data from 1952-2004, forecast the 2008 Electoral College races. The 2008 round of the ANES surveyed respondents in 34 states, including all the important swing states except Missouri and New Hampshire. Table 5 shows that the expectation question provides a much more accurate and informative forecast of 2008 than the intention question. The expectation-based forecasts have smaller absolute and squared errors, and a higher correlation with the actual vote shares. These differences, while meaningful, are not statistically significant. The expectations-based forecast was also closer to the final outcome slightly more often. In the encompassing regression, the expectations-based forecast is statistically significant, while the intentions data are not, and so we cannot reject the null that the intentions data provide no useful information. Finally, as in the full sample, the optimally-weighted average again puts most of the weight on the expectations-based forecast.

Table 5: Comparing the Accuracy of Efficient Forecasts for the 2008 Electoral College

Forecast of Vote Share:	**Efficient Voter Intention:** $E[v_r \mid \widehat{v_r}]$	**Efficient Voter Expectation:** $E[v_r \mid \widehat{x_r}]$	**Test of Equality**
Root Mean Squared Error	0.094 (0.022)	0.085 (0.022)	$t_{33}=1.27$ (p=0.21)
Mean Absolute Error	0.063 (0.012)	0.056 (0.011)	$t_{33}=0.92$ (p<0.3656)
How often is forecast closer?	47.1% (8.7)	52.9% (8.7)	$t_{33}=0.34$ (p<0.7371)
Correlation	61.6%	69.2%	
Encompassing regression: $v_r = \alpha + \beta_v Intention_r + \beta_x Expectation_r$	0.330 (0.291)	0.684*** (0.250)	
Optimal weights: $v_r = \beta Intention_r + (1-\beta) Expectation_r$	24.7% (26.7)	75.3%*** (26.7)	

Notes: ***, **, and * denote statistically significant coefficients at the 1%, 5%, and 10%, respectively. (Standard errors in parentheses). These are assessments of forecasts of the vote share in n=34 elections (the 34 Electoral College races chosen by the ANES). Comparisons in the third column test the equality of the measures in the first two columns. In the encompassing regression, the constant $\hat{\alpha} = 0.032$ (se=0.097).

We now turn to testing the relative accuracy of our two types of polling questions using data from an independent secondary dataset.

VI. Results from Archival Research

Using a variety of archival resources, we have compiled a dataset of any poll from around the world that we could find that asks both an intent and expectation question. The polls are a hodgepodge of USA and non-USA elections, executive and legislative offices, and the full range of national elections to smaller districts. Most of these polls are based on relatively large sample (typically of around 1,000 respondents).

Table 6 summarizes the forecast performance of our two questions in forecasting the winning candidate, mimicking the analysis in Table 2. Again, we use a very coarse performance metric, simply scoring the proportion of races in which the candidate who won a majority in the relevant poll ultimately won the election. The expectation question is correct more often than the intent question in both 0 to 90 and 90 to 180 days before the election; this difference is statistically significant. Indeed this difference is largest in the 90 to 180 day segment. For reasons we are not quite sure of, the expectation question is essentially random after 180 days, in the data in our dataset, but many of those polls are actually a full year and beyond before the election.

Table 6: Comparing the Accuracy from Secondary Dataset

Proportion of observations where the winning candidate was correctly predicted by a majority of respondents to:

	Days Before the Election ≤ 90				90 < Days Before the Election ≤ 180				Days Before the Election > 180			
	Expect	Intent	# Obs	# Elections	Expect	Intent	# Obs	# Elections	Expect	Intent	# Obs	# Elections
President	89.4%	80.7%	161	19	69.2%	61.5%	39	12	59.6%	57.7%	52	11
1936 Electoral College	72.3%	80.9%	47	47	-	-	0	0	-	-	0	0
Governor	78.9%	78.9%	19	9	83.3%	50.0%	6	6	100.0%	100.0%	2	1
Senator	81.8%	90.9%	11	7	-	-	0	0	-	-	0	0
Mayor	100.0%	100.0%	4	2	100.0%	66.7%	3	1	-	-	0	0
Other	80.0%	80.0%	10	9	100.0%	66.7%	3	2	50.0%	50.0%	2	2
USA Total	84.9%	81.3%	252	93	74.5%	60.8%	51	21	60.7%	58.9%	56	14
AUS (Parliament)	88.9%	41.7%	36	3	66.7%	33.3%	21	3	24.4%	66.3%	86	2
GBR (Parliament)	85.0%	90.0%	20	9	100.0%	92.3%	13	7	69.4%	62.9%	62	9
FRA (President)	60.9%	56.5%	23	4	40.0%	20.0%	5	3	-	-	0	0
Other	71.4%	71.4%	7	6	0.0%	0.0%	1	1	0.0%	0.0%	1	1
non-USA Total	79.1%	59.3%	86	22	72.5%	50.0%	40	14	43.0%	64.4%	149	12
Total	83.4%	75.7%	338	115	73.6%	56.0%	91	35	47.8%	62.9%	205	26
Diff (standard error)	7.7%* (2.2)				17.7%* (4.8)				-15.1%* (4.4)			

VII. A Structural Interpretation

At this point, it's worth reflecting on just why it is that the expectation-based forecasts are more accurate. Our working intuition is simply that each of us possesses substantial information that is relevant to forecasting the election outcome, but the voter intention question only elicits part of this information set. By asking about each respondent's expectations, a poll can effectively aggregate this broader information set. We can formalize this idea with a very simple structural model that also does surprisingly well at matching other key facts about voter expectations. We'll begin with a very stripped-down version of the model, and then add greater richness.

A Simple Model

Initially, we conceptualize a survey respondent's expectation as deriving from her knowledge of her own voting intention, as well as those of $m-1$ of her friends, family, and coworkers, effectively creating a survey of m likely voters. We denote the proportion of person i's sample who plan to vote for the Democrat as $\widehat{v_r^i}$. If each person's individual "informal poll" is drawn from an unbiased sample, then $\widehat{v_r^i}$ is a random variable which follows a binomial distribution with mean v_r and variance $v_r(1-v_r)/m$:

$$x_r^{*i} = E^i\left[v_r \middle| \widehat{v_r^i}\right] = \widehat{v_r^i} \sim B\left(v_r, \frac{v_r(1-v_r)}{m}\right) \approx N\left(v_r, \frac{1}{4m}\right) \qquad [18]$$

and hence $E\left[v_r \middle| \widehat{v_r^i}\right]$ approximately follows a normal distribution with mean v_r and variance $\frac{1}{4m}$ (which follows because in close elections, $v_r(1-v_r) \approx \frac{1}{4}$). This suggests that the probability an individual respondent expects the Democrat to win is:

$$E\left[x_r^i \middle| v_r\right] = Prob\left(E\left[v_r \middle| \widehat{v_r^i}\right] > 0.5\right) \approx \Phi\left(2\sqrt{m}(v_r - 0.5)\right) \qquad [19]$$

where $\Phi(.)$ is the standard normal cdf.

Thus, equation [19] suggests that in a simple probit regression of whether an individual forecasts the Democrat to win, the coefficient on the winning (or negative losing) margin of the Democrat candidate reveals $\sqrt{m}$.[13] The intuition is simply that the larger the social circle that each person polls, the more informed they are, and hence the stronger will be the relationship between their expectations and actual outcomes, yielding a higher probit coefficient.

In fact, we have already run the regression described in equation [19]—it is identical to that described in equation [3], which we used to estimate $\widehat{\sigma_e}$. Comparing these equations we see that $2\sqrt{m} = \frac{1}{\sigma_\epsilon}$, or $m = \frac{1}{4\sigma_\epsilon^2}$. Thus given our estimate of $\widehat{\sigma_\epsilon} = 0.150$, we infer that $\hat{m} = 11$ (the exact coefficient is 11.11, and the standard error clustering by state-year and applying the delta method is 1.12).

[13] We can avoid the approximation noted above, and run a probit regression where the dependent variable is instead $\frac{v_r - 0.5}{\sqrt{v_r(1-v_r)}}$ and the coefficient reveals $\sqrt{m}$. This yields similar results.

In the sections below, we shall add more richness to this simple model. And so rather than taking these initial findings too literally, we suggest a somewhat more nuanced interpretation: each poll respondent bases her expectation on an information set that is as rich as would be obtained if she had polled herself, plus a representative sample of ten friends. As we shall see, there are many ways she may have gathered this information.

When social circles have a known partisan bias

Of course, it is unlikely that the social network of any survey respondent is a representative random sample. Instead, people within a social network are more likely to have somewhat similar political attitudes, and hence voting intentions. To investigate the implications, we replace the assumption that the probability of each person within each social circle voting for the Democrat is v_r (as it would be with independent sampling), with the assumption that it is $v_r + \theta_r^{s_i}$ where $\theta_r^{s_i}$ is the bias specific to social circle s to which respondent i belongs.

We begin by considering the case when the respondent knows the partisan bias of their social circle. Given this knowledge, they can effectively de-bias, and as we shall see, the bias has few implications. In particular, given knowledge of this bias, a respondent's efficient expectation is:

$$x_r^{*i} = E^i\left[v_r \middle| \widehat{v_r^i}; \theta_r^{s_i}\right] = \widehat{v_r^i} - \theta_r^{s_i} \sim B\left(v_r, \frac{\left(v_r+\theta_r^{s_i}\right)\left(1-v_r-\theta_r^{s_i}\right)}{m}\right) \approx N\left(v_r, \frac{1}{4m}\right) \qquad [20]$$

where the approximation follows whenever $\theta_r^{s_i}$ is sufficiently small that $(v_r + \theta_r^{s_i})(1 - v_r - \theta_r^{s_i}) \approx 1/4$, and as before, we approximate the binomial distribution with the normal. Notice that this means that when social circles have a known partisan bias, they have a similar mean and variance (in equation [20]) as in the random sampling case (equation [18]). That is, if respondents are aware of the partisan bias in their social circles, then they can effectively de-bias the signal they receive, and equation [19] still (approximately) holds. Thus if people de-bias effectively, none of our conclusions are changed if there is a known partisan bias common to their social circles.

When social circles have unknown correlated shocks

Alternatively, consider the case where respondents do not observe the bias of their social circle. This might occur either because the bias of one's friends is unknown, or alternatively, even if the "typical" bias is known, each social circle is subject to a different shock specific to this election. That is, while each respondent is not aware of the specific bias affecting their social circle, they are aware that members of their social circle likely have correlated preferences, parameterized by an unobserved shock common to their social circle, $\eta_r^{s_i} \sim N(0, \sigma_\eta^2)$. This means that the result of a respondent i's informal poll is a random variable drawn from a binomial distribution with mean $v_r + \eta_r^{s_i}$. For the reasons described above, we approximate this by a normal distribution with mean $\eta_r^{s_i}$ and variance $\frac{1}{4m}$. Thus, across the whole population, the different signals people get are, on average, unbiased, but with total variance of $\sigma_\eta^2 + \frac{1}{4m}$, which we rewrite as $\frac{1}{4m(1-\lambda)}$ (where $\lambda = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \frac{1}{4m}}$ is the intra-cluster correlation—the share of the variance that is due to the common shock).

Because survey respondents don't observe the common shock $\eta_r^{s_i}$, they can't use it to de-bias. As such, their expectations are simply equal to the mean of their informal poll:

$$x_r^{*i} = E^i\left[v_r \middle| \widehat{v_r^i}\right] = \widehat{v_r^i} \sim N\left(v_r, \frac{1}{4m(1-\lambda)}\right) \quad [21]$$

The key difference from the earlier case is the variance term is larger which reflects the fact that sampling from a sub-population with correlated voting intentions is less informative, relative to random sampling. Given this less informative signal, we replace equation [19] with the following expression for the probability that a poll respondent expects the Democrat to win:

$$E\left[x_r^i \middle| v_r\right] = Prob\left(E\left[v_r \middle| \widehat{v_r^i}\right] > 0.5\right) \approx \Phi\left(\left[2\sqrt{m(1-\lambda)}\right](v_r - 0.5)\right) \quad [22]$$

Thus the coefficient from estimating equation [3] reveals the term in square brackets, thereby yielding only partial identification of m. For instance, if $\lambda = 0$, we get the same expression as above: $m = \frac{1}{4\sigma_\epsilon^2}$, which combined with our estimate of $\widehat{\sigma_\epsilon} = 0.150$, implies that that $\widehat{m} = 11$. If there is substantial correlation in the views of your social circle (say, half the variation in people's signals is due to their group affiliations so that $\lambda = 0.5$), then $m = 22$. We now turn to other data on social networks to identify λ, which allows for point identification of m.

The 2000 National Election Studies post-election survey asked "From time to time, people discuss government, elections and politics with other people. I'd like to ask you about the people with whom you discuss these matters. These people might or might not be relatives. Can you think of anyone?" The survey then allowed respondents to name up to four members of their social circle, and for each named person, they were asked: "How do you think [name] voted in the election? Do you think he/she voted for Al Gore, George Bush, some other candidate, or do you think [name] didn't vote?" We analyze the correlation of these responses with how the respondent reported voting, in order to estimate

λ.[14] In particular we stack all respondents, plus the members of their social circles into one dataset, and estimate the following random effects model:

$$I\left(v_r^i = 1\right) = r_r + \eta_r^{s_i} + \zeta_r^i \qquad [23]$$

where the dependent variable describes whether individual i voted for the Democrat, and this is a function of r_r which is a (state×year) election race fixed effect, $\eta_r^{s_i}$ which is the random effect common to respondent i's social circle, and ζ_r^i is the individual-level error term. Our interest is in estimating the share of the variation in voting patterns within each race that is common to people's social circles, $\lambda = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\zeta^2}$. Estimating this random effects model yields an estimate of $\lambda = 0.45$, suggesting that slightly less than half of the variation in voting patterns is common to social circles.[15] Plugging this back into [22] yields a point estimate for $\hat{m}$ =20.

The point of this slightly more general model is to qualify our earlier observation. We had suggested that when people answer the voter expectation question they draw upon personal knowledge whose information content is similar to that of a random sample with an effective sample size of eleven. This somewhat more general model suggests that this could also reflect them drawing on a larger sample that is less informative because observations aren't independent—the information content of a random sample of m respondents is equivalent to a sample of $m' = \frac{m}{1-\lambda}$ drawn from a sample with intra-cluster correlation coefficient λ. These two interpretations are observationally equivalent in our main dataset, and none of our empirical observations about the relative forecast accuracy of the two survey questions depends on which is correct.

VIII. What information is being aggregated?

Let's take stock. At this point we've documented that data on voter expectations yield more accurate forecasts than data on voter intentions. Presumably this reflects the voter expectations question yielding more information. We now turn to evaluating what this extra information is. One possibility is that the voter expectations question extracts idiosyncratic information—such as the voting intentions of

[14] We analyze the responses of only those respondents or members of their social circle who plan to vote for one of the two major parties.

[15] Note that we are estimating λ from the correlation in people's perceptions of how their networks actually voted, but we are using this to speak to the correlation of their voting intentions prior to the election.

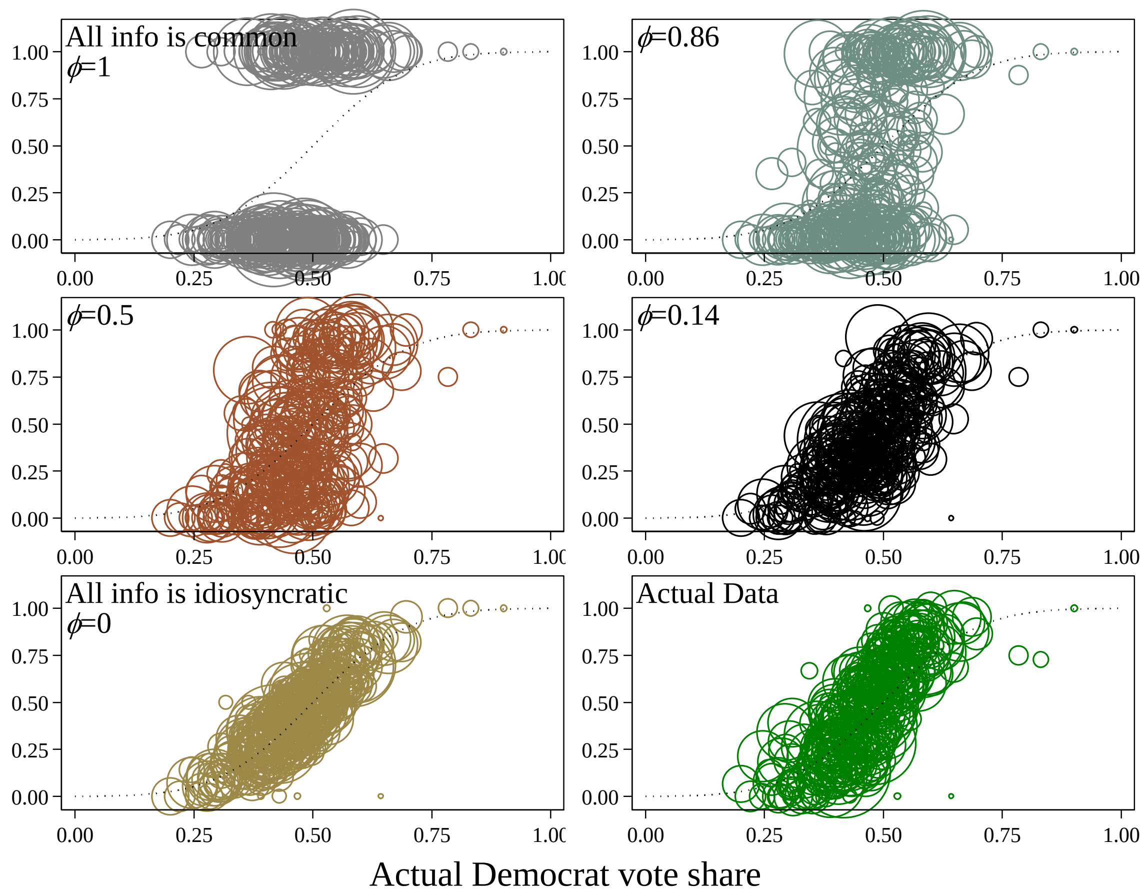
friends and family, or perhaps the success of the respective campaigns in very local communities. Indeed, this has informed our modeling so far, and based on this conjecture, we estimated that each person draws on a sample of $m = 20$ people. Another possibility is that respondents are using common information, such as newspaper articles, public opinion polls, televised debates or past election results to inform their expectations. If this were true, then the true number of friends each person draws on would be zero. More importantly, from a forecasting perspective observing many sources of idiosyncratic information is more valuable than simply observing common information many times. We now turn to enriching the model to allow for common information so that we can evaluate the extent to which the extra information being aggregated is idiosyncratic, versus common information. To preview: We find that the vast majority of the information is idiosyncratic information, such as that obtained from an informal poll of friends.

The Extent of Disagreement

We begin by considering two extreme cases. If all information that survey respondents use in assessing their expectations is common information—for instance, if people read the same newspaper and this is the only information shaping their expectations—then survey respondents will never disagree. That is, if the common signal is pro-Democrat, they will all pick the Democrat to win; if it is pro-Republican, they will all pick the Republican. By contrast, if all information is idiosyncratic to each respondent's social circles and local community, then different draws of idiosyncratic information will lead people to have different expectations. Moreover, if these idiosyncratic signals are actually informative, the highest levels of disagreement will be occur in the closest elections. And intermediate cases will yield intermediate degrees of disagreement.

In order to see this intuition at work, we ran several simulations. For each person, we drew an independent unbiased idiosyncratic signal of the election outcome, and for each election, an independent unbiased common signal. In the "common shocks" simulation, each respondent only puts weight on the common signal. Then, for each election, we add up the proportion who expects the Democrat to win, and plot this against the actual Democrat vote share in the top-left panel of Figure 7. Notice that there is no disagreement in this figure—each plot shows that either all or none of the respondents expect the Democrat to win. At the polar opposite case, each person puts all the weight on the idiosyncratic signal. The results—shown in the bottom-left panel of Figure 7—reveal substantially more disagreement. In between, we varied the weights so that the proportion of the variance in people's information set that is due to common information, ϕ, took on intermediate values. In each simulation, the precision of the respondent's expectations is set to match that in the data (that is, $\sigma_\epsilon = 0.150$), and so the weight on the unit variance common signal is equal to $\sqrt{\phi\sigma_\epsilon^2}$ and that on the unit variance idiosyncratic signal is $\sqrt{(1-\phi)\sigma_\epsilon^2}$.) As can be seen, the higher the share of information that is common (that is, higher values of ϕ), the more likely the share of the population who expect a candidate to win is close to zero or one; more frequent intermediate values occur only with a lower value of ϕ.

Figure 7: The Influence of Common Information on the Extent of Disagreement



The bottom right panel shows the actual data. As can be seen, this closely resembles the simulation for the case $\phi = 0.14$. As we'll see in a moment, this is no accident.

A Random Effects Probit Model

In order to make progress, we need to add a bit more structure. As before, each person receives an unbiased idiosyncratic signal, $\widehat{v_r^i}$, and as we derived in equation [21], $\widehat{v_r^i} \sim N(v_r, \frac{1}{4m(1-\lambda)})$. All this is familiar. But we now add the fact that all voters in an election also observe an unbiased common signal, $c_r \sim N(v_r, \sigma_c^2)$, which is orthogonal to the idiosyncratic shock: $E\left[\widehat{v_r^i} c_r\right] = 0$. Here, c_r reflects any information common across respondents—newspapers, polls, or perhaps a common prior. If people use this information efficiently, their (latent) expectation of each candidate's vote share, x_r^{*i} is a precision-weighted average of each signal:

$$x_r^{*i} = E^i[v_r|\widehat{v_r^i}; c_r] = \frac{\sigma_c^{-2}c_r + 4m(1-\lambda)\widehat{v_r^i}}{\sigma_c^{-2}+4m(1-\lambda)} = v_r + \underbrace{\frac{\sigma_c^{-2}}{\sigma_c^{-2}+4m(1-\lambda)}(c_r - v_r)}_{common} + \underbrace{\frac{4m(1-\lambda)}{\sigma_c^{-2}+4m(1-\lambda)}(\widehat{v_r^i} - v_r)}_{idiosyncratic} \quad [24]$$

Thus, equation [24] expands upon equation [1] for the case in which there are both common and idiosyncratic errors, and the sum of the common plus idiosyncratic error terms, $\epsilon \sim N(0, \frac{1}{\sigma_c^{-2}+4m(1-\lambda)})$. Consequently, this equation, plus the choice model outlined in equation [2] together imply that we can estimate the relative importance of these components by estimating a random effects probit model, where the random effect is the information common to all respondents in a particular election. Thus, in the equation below, we explain whether each respondent expected a candidate to win, with a variable describing the extent to which the vote share garnered by that candidate exceeds the 50% required to win, and we allow for both idiosyncratic and common shocks across respondents.

$$E[x_r^i|v_r] = E[x_r^{*i} > 0.5|v_r] = \Phi\left(\frac{v_r - 0.5}{\sqrt{\frac{1}{\sigma_c^{-2}+4m(1-\lambda)}}}\right) = \Phi\left(\sqrt{\sigma_c^{-2} + 4m(1-\lambda)}(v_r - 0.5)\right) \quad [25]$$

Our interest in running this model is in estimating the share of the variance contributed by the common random effect, which we denote ϕ. The intuition for what is identifying this can be seen in Figure 7. Estimating this random effects probit model by maximum likelihood yields an estimate of $\hat{\phi} = 0.14$ (with a standard error of 0.015). That is, the model suggests that very little of the variance in people's expectations—across elections and respondents—is common to respondents within an electoral race.

This same model also identifies the underlying structural parameters. To see this, note that we estimate two moments—the share of variance accounted for by the common effect, ϕ and the estimated regression coefficient—and these are (non-linear) functions of two parameters, σ_c^2 and $m(1-\lambda)$.[16] We

[16] Standard regression packages use a slightly different normalization than that written in equation [25], normalizing so that the idiosyncratic error term $\frac{4m(1-\lambda)}{\sigma_c^{-2}+4m(1-\lambda)}(\widehat{v_r^i} - v_r)$ has a variance of one. To follow this normalization, we rewrite equation [25] as follows:

$$E[x_r^i|v_r] = p\left(\underbrace{\frac{\sigma_c^{-2}+4m(1-\lambda)}{2\sqrt{m(1-\lambda)}}}_{\beta}\left(v_r + \frac{\sigma_c^{-2}}{\sigma_c^{-2}+4m(1-\lambda)}(c_r - v_r) + \frac{4m(1-\lambda)}{\sigma_c^{-2}+4m(1-\lambda)}(\widehat{v_r^i} - v_r)\right) > 0.5\right).$$

estimate $\widehat{\sigma_c^2} = 0.17$ (se 0.023), which suggests that the common signal (or common prior) is extremely uninformative. We also estimate $m\widehat{(1-\lambda)} = 9.0$ (se=0.84). Using our earlier estimate of $\hat{\lambda} = 0.45$, this yields an estimate of $\hat{m} = 16.2$. To put this into perspective, our previous estimates of the size of each person's social circles m was based on the assumption that this was their only source of information—that $\phi = 0$. But because this common signal is so uninformative, it only changed our estimate slightly.

IX. The Correlation Between Voter Expectations and Intentions

While our investigation of data on voter expectations as inputs to forecasting models is largely novel, these data have long been of interest to political psychologists. The earliest such finding is Hayes (1936) who found that a majority of those intending to vote for Hoover in the 1932 election expected him to win, even as nearly all of those voting for Roosevelt expected him to win, a result that he inferred (p.190) shows that "the expectations of voters are not merely reasoned judgments based on the best factual data obtainable." Subsequently, Lazarsfeld, Berelson and Gaudet's (1944) study of voters in Erie County in 1940 found "a close relationship between vote intention and expectation of winner." For the 1952-80 election cycles, Granberg and Brent (1983) document a strong correlation between which candidate a respondent intends to vote for, and who they expect to win. This has been interpreted as evidence of "wishful thinking… in which people are motivated to alter their expectations to fit their desires" (Granberg and Holmberg, 1988, p.138). Krizan, Miller and Johar (2011, p.140) call the link between preferences and expectations "one of the most robust findings in social psychology." Alternatively, this correlation has been interpreted as evidence that "people are reliant or dependent on others for cues in the development of preferences and as guides to behavior," which can yield bandwagon effects as voters prefer to vote for the winning candidate (Granberg and Holmberg, 1988, p138). However, our simple model provides a rational re-interpretation.

Table 7 shows the link between individual voter intentions and expectations in our data. Across these election, 68.8% of those intending to vote for the Democrat expect the Democrat to win, just as 73.0% of those intending to vote for the Republican expect the Republican to win. All told, 70.9% of survey respondents expect the candidate they intend to vote for to also win their state, and the correlation between voter intentions and expectations is 0.42.

As such, the estimated regression coefficient on $\beta = \frac{2\sqrt{m(1-\lambda)}}{\sigma_c^{-2}+4m(1-\lambda)}$ and the intra-cluster correlation coefficient $\phi = \frac{\sigma_c^{-2}}{\sigma_c^{-2}+4m(1-\lambda)}$. Solving for the structural parameters yields $\widehat{\sigma_c^2} = \frac{1}{\hat{\beta}^2\hat{\phi}(1-\hat{\phi})}$ and $\widehat{m}(1-\lambda) = \frac{\hat{\beta}^2(1-\hat{\phi})^2}{4}$.

Table 7: Relationship Between Individual Voters' Intentions and Their Expectations

Panel A: Cross-Tabs

	Expect Democrat to win this state	Expect Republican to win this state
Intend to vote Democrat	33.9%	15.4%
	68.8% of intending Democrats	31.2% of intending Democrats
Intend to vote Republican	13.7%	37.1%
	27.0% of intending Republicans	73.0% of intending Republicans

Panel B: Summary Statistics

Proportion who expect their candidate to win	70.9%
Correlation between intentions and expectations	0.42

However, in our model, much of this correlation is easily explained. The basic logic is that each respondent relies on an informal poll of *themselves* plus their friends. As such, some correlation is to be expected between one's own voting intentions, and one's expectations. The remaining issue is how large this correlation will be.

In order to highlight best highlight the intuition, we begin by evaluating the simple version of our model in which respondents' expectations are based only on a representative poll of themselves plus $m-1$ friends. We denote of the proportion of your friends—excluding you—who intend to vote for the Democrat $\widehat{v_r^{-\iota}}$. Given our earlier assumptions, $\widehat{v_r^{-\iota}} \sim N(v_r, \frac{v_r(1-v_r)}{m-1})$. If you include your own voting your voting intentions in your informal poll, then you will expect the Democrat to win if $1+(m-1)\widehat{v_r^{-\iota}} > \frac{m}{2}$, which simplifies to: $\widehat{v_r^{-\iota}} > \frac{\frac{m}{2}-1}{m-1}$. Hence conditional on intending to vote for the Democrat, the probability that a survey respondent expects the Democrat to win is:

$$\begin{aligned} Prob\left(x_r^i = 1 | v_r^i = 1; \widehat{v_r^{-\iota}}\right) &= Prob\left(1+(m-1)\widehat{v_r^{-\iota}} > \frac{m}{2}\right) = Prob\left(\widehat{v_r^{-\iota}} > \frac{\frac{m}{2}-1}{m-1}\right) \\ &\approx \Phi\left(2\sqrt{m-1}\left(v_r - 0.5 + \frac{1}{2(m-1)}\right)\right) \approx \Phi(6.4(v_r - 0.45)) \end{aligned} \quad [26]$$

where the approximation comes from $v_r(1-v_r) \approx \frac{1}{4}$ in close elections, from the normal approximation to the binomial, and from substituting in the estimate of $\hat{m} = 11.1$ which applies to this simpler form of the model. And symmetrically, if you intend to vote Republican, then you expect the Democrat to win only if $0+(m-1)\widehat{v_r^{-\iota}} > \frac{m}{2}$. Hence the probability that a survey respondent who intends to vote Republican expects the Democrat to win is:

$$Prob\left(x_r^i = 1 | v_r^i = 0; \widehat{v_r^{-\iota}}\right) = Prob\left(0 + (m-1)\widehat{v_r^{-\iota}} > \frac{m}{2}\right) = Prob\left(\widehat{v_r^{-\iota}} > \frac{\frac{m}{2}}{m-1}\right)$$
$$\approx \Phi\left(2\sqrt{m-1}\left(v_r - 0.5 - \frac{1}{2(m-1)}\right)\right) \approx \Phi(6.4(v_r - 0.55)) \quad [27]$$

Thus in a close election (when $v_r = 0.5$), the simple fact that you are in your own information set used in forming your expectations is sufficient to lead 62.6% of people to say that they expect the election to be won by the candidate they intend to vote for. In turn, this model implies that the correlation between voters' intentions and expectations will be $Correlation\left(v_r^i, x_r^i\right) \approx \frac{\Phi(6.4(v_r - 0.45)) - \Phi(6.4(v_r - 0.55))}{2\sqrt{\Phi(6.4(v_r - 0.55))(1 - \Phi(6.4(v_r - 0.55)))}}$, which in a close election yields a correlation of 0.25.[17] In more lop-sided races (say, $v_r = 0.6$), these biases remain important, and 82.9% of those voting for the Democrat expect the Democrat to win, compared with 62.6% of those voting for the Republican, yielding a correlation between voter intentions and expectations of 0.24. Thus, our simple (rational) model delivers a quantitatively important correlation between voters' intentions and expectations, albeit one somewhat smaller than observed in our data.

We now turn to assessing the richer version of our model. To preview, the math gets a little more involved, but we get roughly similar results, although the model explains slightly less of the correlation between intentions and expectations. In the full version of the model we allow for both idiosyncratic and common sources of information, and for correlated shocks within the social networks of each respondent. Consequently, the process by which each person forms their expectations is described by equation [24], which in turn implies that a respondent will say they expect the Democrat to win if $v_r + \phi(c_r - v_r) + (1-\phi)\left(\frac{1+(m-1)\widehat{v_r^{-\iota}}}{m} - v_r\right) > 0$. Thus, the probability that someone who intends to vote for a Democrat expects the Democrat to win is:

$$Prob\left(x_r^i = 1 | v_r^i = 1; \widehat{v_r^{-\iota}}, c_r\right) = Prob\left(v_r + \phi(c_r - v_r) + (1-\phi)\left(\frac{1+(m-1)\widehat{v_r^{-\iota}}}{m} - v_r\right) > 0\right)$$
$$\approx \Phi\left(2\frac{\sqrt{(m-1)(1-\lambda)}}{(1-\phi)}\frac{m}{m-1}\left(v_r - 0.5 + \phi(c_r - v_r) + \frac{1-\phi}{m}(1 - v_r)\right)\right) \quad [28]$$
$$\approx \Phi(7.18(v_r - 0.5 + 0.14(c_r - v_r) + 0.05(1 - v_r))$$

[17] This follows from the fact that the correlation of two binary variables is equal to: $\frac{p(x_r^i=1, v_r^i=1)p(x_r^i=0, v_r^i=0) - p(x_r^i=1, v_r^i=0)p(x_r^i=0, v_r^i=1)}{\sqrt{p(x_r^i=0)p(x_r^i=1)p(v_r^i=0)p(v_r^i=1)}}$

Notice that this expression conditions on the unobservable common information. We could integrate over possible values of c_r to find the unconditional expectation, but because the expression in equation [28] is approximately symmetric for close elections, we get the following simple approximation:[18]

$$Prob\left(x_r^i = 1 \middle| v_r^i = 1; \widehat{v_r^{-\iota}}\right) \approx \Phi\left(2\frac{\sqrt{(m-1)(1-\lambda)}}{(1-\phi)}\frac{m}{m-1}\left(v_r - 0.5 + \frac{1-\phi}{m}(1-v_r)\right)\right)$$

$$\approx \Phi\big(7.58(v_r - 0.47)\big) \qquad [29]$$

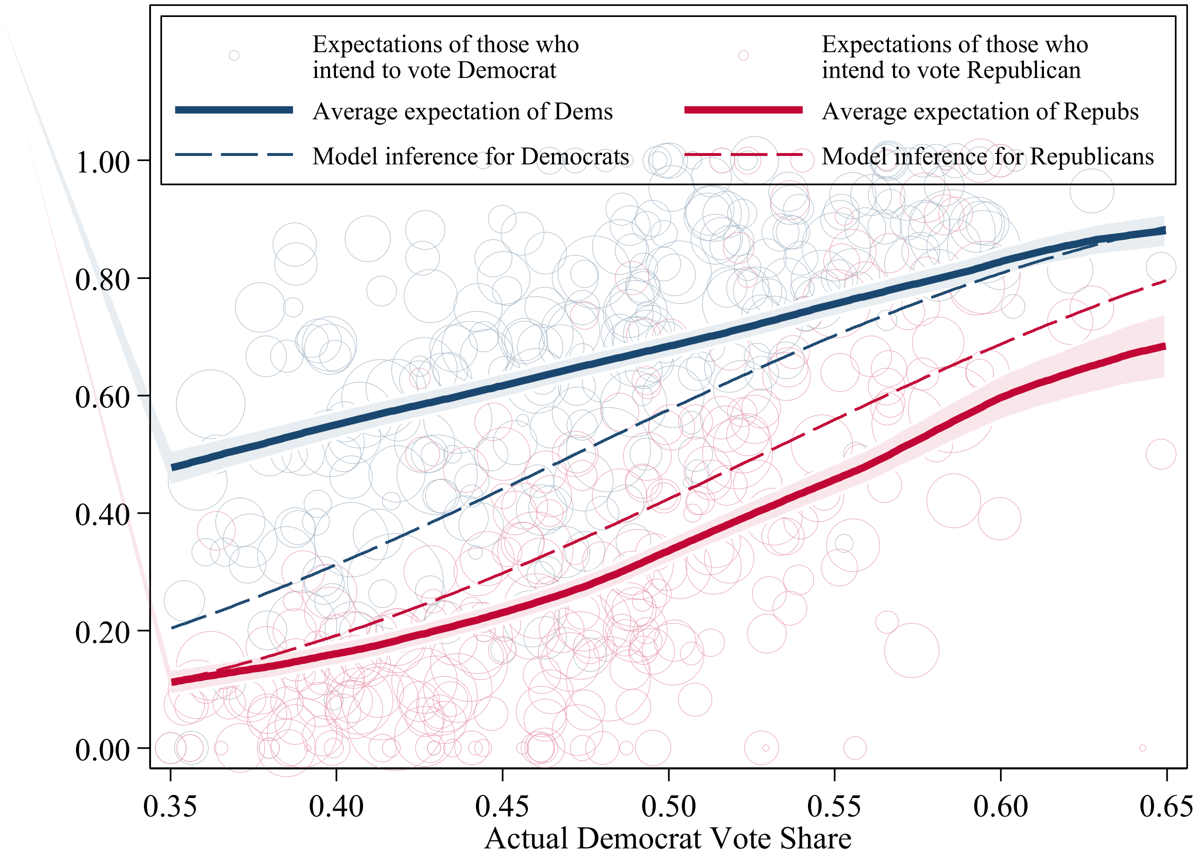
where the precise values come from substituting in our estimates of $\widehat{m} = 16.2$, $\hat{\lambda} = 0.45$, and $\hat{\phi} = 0.14$. Symmetrically, the probability that someone intending to vote for the Republican expects the Democrat to win is: $Prob\left(x_r^i = 1 \middle| v_r^i = 0; \widehat{v_r^{-\iota}}\right) \approx \Phi(7.58(v_r - 0.53))$. These results are quite similar to the simple model shown in equations [26] and [27]. Their implications are similar, too: For a close election, this full version of the model implies that 57.6% of voters will say that they expect "their" candidate to win, and hence that the correlation between individual voters' expectations and their intentions will be 0.15.

Figure 8 summarizes both the data and our model inferences more fully. Each circle shows the proportion of poll respondents who expect the Democrat to win, with separate circles shown for the respondents who intend to vote for the Democrat (in blue) and those who intend to vote for the Republican (in red). The fact that intentions are correlated with expectations leads the former to lie above the latter. We summarize these patterns with a local running mean, shown with a 95% confidence interval. Comparing these two lines, we see that Democrats are typically about 40 percentage points more likely to expect their candidate to win. The inferences from the full model are also shown as dashed lines, and these suggest that our model can explain about half of this difference.

---

[18] That is, $Prob(x_r^i = 1 \middle| v_r^i = 1; \widehat{v_r^{-\iota}}) = \int \Phi\left(2\frac{\sqrt{(m-1)(1-\lambda)}}{(1-\phi)}\frac{m}{m-1}\left(v_r - 0.5 + \phi(c' - v_r) + \frac{1-\phi}{m}(1-v_r)\right)\right)dF(c')$. When $v_r = 0.5$, the integrand is symmetric around v_r, and hence this is equal to $\Phi\left(2\frac{\sqrt{(m-1)(1-\lambda)}}{(1-\phi)}\frac{m}{m-1}\left(v_r - 0.5 + \frac{1-\phi}{m}(1-v_r)\right)\right)$. For close elections, the same argument holds as an approximation.

Figure 8: Proportion Expecting the Democrat to Win, by Whether Intending to Vote Democrat or Republican



Thus, our simple (rational) model delivers a quantitatively important correlation between voters' intentions and expectations, albeit one somewhat smaller than observed in our data. This finding should be contrasted with the claim by Granberg and Holmberg (1988, p.139) that "under the rational democratic model, even when there is sufficient variation in preferences and expectations, there would be no systematic relationship between preference and expectation." Indeed, they go on to claim that "a significant relationship between preference and expectation can indicate a departure from rationality." By contrast, in our simple model, this correlation is due to voters making rational inferences from limited data. Interestingly, this finding was hinted at by Hayes (1936, p.190), when he noted that the correlation between expectations and intentions may reflect either "a real case of wish-thinking, or merely a case of over-subjection to a strictly limited mass of data."

X. Efficient Poll-Based Forecasts from Non-Random Samples

Under many circumstances we are left making forecasts based off polls with overtly non-random samples. An extreme example of that would be a poll where all of the respondents support just one of the candidates. In this section we test the accuracy of forecasts based off of the expectation question data originating from respondents who are exclusively voting for one major party or the other, by comparing those forecasts to forecasts created from the full sample of intentions. For the non-random samples, the sample size drops from an average of 38 to about 19 respondents.

We start by revisiting equation [12] from Section IV: $E[x_r|\widehat{x_r}] = \mu_x + \frac{\sigma_{\bar{x}}^2}{\sigma_{\hat{x}}^2}(\widehat{x_r} - \alpha_x - \mu_x)$, where μ_x is the mean across all elections of the proportion of the population who expect the Democrat to win and σ_x^2 is the corresponding variance. These values are the same in this biased sample as they are in Section IV.[19] $\sigma_{\hat{x}}^2$ is the variance of the sample estimator, and our bias parameter, α_x, will attempt to account for the oversample people who expect Democrats to win (which should be heavily positive in a Democratic only sample and heavily negative in a Republican only sample).

The bias term is $\hat{b} = \sum(\widehat{x_r} - \Phi\left(\frac{v_r - 0.5}{\widehat{\sigma_\epsilon}}\right))/R$. This is 0.203 (0.016) for the Democratic sample and -0.112 (0.012) for the Republican sample. In Section IV, we conclude that the full dataset has a bias of 0.042, approximately half the difference in the absolute bias of the Democratic and Republican supporters.

We make the same assumptions as in Section IV in regard to $\widehat{\sigma_{\hat{x}}^2}$, which, combining equations [13] and [14], is $\widehat{\sigma_x^2} + \frac{(1+(n_r-1)\rho_x)}{4n_r}$. We are going to have to re-derive ρ for these two biased samples, as the clustering is going to have a totally different affect within party. Both samples have the same $\widehat{\sigma_x^2}$ and about half the observations as the full samples. This yields shrinkage estimators (i.e., $\frac{\sigma_{\bar{x}}^2}{\sigma_{\hat{x}}^2}$) for the Democratic sample with an average of 0.59, ranges from 0.14 to 0.74. Not surprisingly, the full dataset has shrinkage estimators that are on average 7.5 percentage points larger, ranging from a low of the same to a higher of 54 percentage points larger. For the Republican sample, the differences are not as stark, with the Republican sample having slightly smaller $\widehat{\rho_x}$ than the full sample; the full sample has shrinkage

[19] There is a slight variation in the variables because 5 of the races in our dataset have no Democratic supporters and 4 no Republican supporters. Thus, when we drop those races from the dataset when we explore forecast accuracy of the respective non-random samples.

estimators that are on average 2.5 percentage points larger, but they range from a low of -7 percentage points smaller to 42 percentage points larger.

In Table 8 we show that forecasts based on the expectations of the Democratic voters only or the Republican voters only, a non-random selection of approximately half of the sample, yield a lower root mean squared error, mean absolute error, and higher correlation than forecasts based on the full sample of voter intention. The first two rows show that the expectation-based forecasts yield both a root mean squared error and mean absolute error that is less than the intention-based forecasts by a statically significant amount for the Republican sample and a weakly significant amount for the Democratic sample. The third row shows that the expectation-based forecasts are also the more accurate forecast in the majority of elections. The expectation-based forecasts are still more highly correlated with actual vote shares than are the intention-based forecasts by a sizable margin. In the Fair-Shiller regression, the expectation-based forecasts have a much larger weight than the intention-based forecast; both are statistically significant, so the intention-based data may be providing some unique information. Finally, an optimally weighted average still puts nearly 60% of the weight on the expectation-based forecast in the Democratic sample and just over 70% in the Republican sample.

Table 8: Comparing the Accuracy of Efficient Forecasts of Vote Shares from Biased Samples

	Democratic Sample		Republican Sample	
Forecast of Vote Share:	Intention: $E[v_r\|\widehat{v_r}]$ Full Sample	Expectation: $E[v_r\|\widehat{x_r}]$ Democratic Supporters	Intention: $E[v_r\|\widehat{v_r}]$ Full Sample	Expectation: $E[v_r\|\widehat{x_r}]$ Republican Supporters
Root Mean Squared Error	0.075 (0.005)	0.070 (0.006)	0.071 (0.004)	0.062 (0.004)
Mean Absolute Error	0.056 (0.003)	0.050 (0.003)	0.054 (0.003)	0.048 (0.002)
How often is forecast closer?	46.7% (2.9)	53.3% (2.9)	44.0% (2.8)	56.0% (2.8)
Correlation	0.592	0.664	0.604	0.718
Encompassing regression: $v_r = \alpha + \beta_v Intention_r + \beta_x Expectation_r$	0.625*** (0.078)	0.790*** (0.071)	0.489*** (0.077)	0.786*** (0.065)
Optimal weights: $v_r = \beta Intention_r + (1-\beta) Expectation_r$	38.5%*** (6.6)	61.5%*** (6.6)	29.8%*** (6.3)	70.2%*** (6.3)
	306 Elections		307 Elections	

Notes: ***, **, and * denote statistically significant coefficients at the 1%, 5%, and 10%, respectively. (Standard errors in parentheses). These are assessments of forecasts of the Democrat's share of the two-party vote. In the encompassing regression with the Democratic sample, the constant $\hat{\alpha} = -0.196$ (se=0.036)*** and in the Republican sample, the constant $\hat{\alpha} = -0.129$ (se=0.031)***. A few of the 311 observations are dropped in each sample, because there is no intention for the given party.

XI. Discussion

A Fox News poll, in the field on September 8 and 9 of 2008, prompted a headline trumpeting McCain's lead in the intent question, but buried Obama's continued lead in the expectation question.[20] This is an editorial choice that almost any polling or news organization would have made. Yet, while McCain and Obama occasionally traded the lead in the intent question, Obama, the eventual winner, led in every expectation question we could find. We hope that this paper will give election analysts a new lens for interpreting polling data, emphasizing data from questions asking voters about their expectations, rather than their intentions. Indeed, our results suggest that expectations-based forecasts are much more powerful predictors of election outcomes.

The structural interpretation of our data helps illustrate why expectations are such a powerful polling tool. The answer swe receive from the expectation question are about as informative as if they were themselves based on a personal poll of approximately twenty friends, family, and coworkers. This "turbocharging" of the effective sample size makes the expectation question remarkably valuable with small sample sizes. Moreover, because our model gives insight into the correlation between voting expectations and intentions, even samples with a strong partisan bias can be used to generate useful forecasts.

The key insight from our study—that analysts pay greater attention to polls of voter forecasts—in fact represents a return to historical practice. In the decades prior to the advent of scientific polling, the standard approach to election forecasting involved both newspapers and business associations writing to correspondents around the country, asking who they thought would win. We are in many respects, recommending a similar practice. Having shown the usefulness of this approach for forecasting elections, we hope that future work will explore how similar questions can be used to provide better forecasts in a variety of market research contexts from forecasting product demand topredicting electoral outcomes, to better measuring consumer confidence.

[20] http://www.foxnews.com/story/0,2933,420361,00.html

XII. References

Berg, Joyce A., Forrest D. Nelson, and Thomas A. Rietz. "Prediction Market Accuracy in the Long Run." *International Journal of Forecasting* 24, no. 2 (2008): 283-298.

Berg, Joyce E., and Thomas A. Rietz. "The Iowa Electronic Market: Stylized Facts and Open Issues." In *Information Markets: A New Way of Making Decisions in the Public and Private Sector*, by Robert W. Hahn and Paul (eds) Tetlock. Washington, D.C.: AEI-Brookings Press, 2006.

Campbell, James E. *The American Campaign.* College Station: Texas A&M University Press, 2000.

DeBell, Matthew. *How to Analyze ANES Survey Data.* ANES TEchnical Report Series, Palo Alto, CA and Ann Arbor, MI: Stanford University and University of Michigan, 2010.

Fair, Ray, and Robert Shiller. "Comparing Information in Forecasts from Econometric Models." *American Economic Review* 80, no. 3 (1990): 375-389.

Fair, Ray, and Robert Shiller. "The Informational Content of ex-Ante Forecasts." *The Review of Economics and Statistics* 71, no. 2 (1989): 325-331.

Granberg, Donald, and Edward Brent. "When Prophecy Bends: The Preference-Expectations Link in U.S. Presidential Elections." *Journal of Personality and Social Psychology* 45, no. 3 (1983): 477-491.

Granberg, Donald, and Soren Holmberg. *The Political System Matters: Social Psychology and Voting Behavior in Sweden and the United States.* Cambridge University Press, 1988.

Hayes, Samuel P. "The Predictive Ability of Voters." *The Journal of Social Psychology* 7, no. 2 (1936): 183-191.

Krizan, Zlatan, Jeffrey C. Miller, and Omesh Johar. "Wishful Thinking in the 2008 U.S. Presidential Election." *Psychological Science* 21, no. 1 (2011): 140-146.

Lazarsfeld, P., B. Berelson, and H. Gaudet. *The people's choice: how the voter makes up his mind in a presidential campaign.* New York: Columbia University Press, 1944.

Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. *The people's choice: How the voter makes up his mind in a presidential campaign.* New York: Columbia University Press, 1944.

Lock, Kari, and Andrew Gelman. "Bayesian Combination of State Polls and Election Forecasts." *Political Analysis* 18, no. 3 (2010): 337-348.

Moulton, Brent. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units." *The Review of Economics and Statistics* 72, no. 2 (1990): 334-338.

Robinson, Claude E. "Recent Developments in the Straw-Poll Field." *Public Opinion Quarterly* 1, no. 3 (1937): 45-56.

Rothschild, David. "Forecasting Elections: Comparing Prediction Markets, Poll and Their Biases." *Public Opinion Quarterly* 73, no. 5 (2009): 895-916.

The American National Election Studies. *Time Series Cumulative Data File, 1948-2008.* Prod. Stanford University and the University of Michigan. August 25, 2011.

Wolfers, Justin, and Eric Zitzewitz. "Prediction Markets." *Journal of Economic Perspectives* 18, no. 2 (2004): 107-126.