International Journal of Forecasting 31 (2015) 952–964

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Combining forecasts for elections: Accurate, relevant, and timely

David Rothschild

Microsoft Research, New York, NY, USA

ARTICLE INFO

Keywords:
Election forecasting
Surveys
Econometric models
Prediction markets
Combining forecasts
Probability forecasting

ABSTRACT

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

1. Introduction

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

The need to transform raw polling data into a forecast is conclusive in the literature. Campbell (2008) clearly illustrates the anti-incumbency bias, whereby incumbents have lower polling values than the actual election results, and the fading of early leads in polls, whereby election results are tighter than polling numbers. Erikson and Wlezien (2008a) show that translating raw polling data into a forecast makes it more accurate for both the estimated vote share and the probability of victory. Rothschild (2009) improves on the work of Erikson and Wlezien (2008a) by aggregating the daily polls over time, eliminating noisy daily fluctuations, then translating them into a forecast. At the same time, Rothschild (2009) designed his poll-based forecast to be the most accurate forecast using the same general model as Erikson and Wlezien (2008a), leaving open the possibility of creating even more accurate

E-mail address: DavidMR@Microsoft.com.

http://dx.doi.org/10.1016/j.ijforecast.2014.08.006

transformations through more advanced models of the aggregation and subsequent translation of the polling data into forecasts. The most recent advances in creating forecasts from polls have been in the area of aggregation, including eliminating poll company specific effects and combining the snapshot for any given state with other state and national polls. In this paper, when available, we use both the most transparent and the most efficient method possible, without these further steps, but with Stanford's Simon Jackman's interpretation of these steps, as made available through Pollster.com (Jackman, 2005).

There is a massive body of literature on the modeling of fundamental data, which has found that most models are not useful as forecasts, but rather explain the correlations between different variables and election outcomes. These models use a variety of economic and political indicators, such as past election results, incumbency, presidential approval ratings, economic indicators, ideological indicators, biographical information, policy indices, military situations, and facial features of the candidates. Hummel and Rothschild (2013) provide a substantial list of such models; however, there are several reasons why they are generally not useful for producing forecasts. First, many models are difficult to duplicate, such as that of Armstrong, Green, Jones, and Wright (2010), which utilizes pictures of the candidates. Second, many models incorporate pre-election polls or other late-arriving data; for example, Lock and Gelman (2010) use a model that cannot be resolved until October of the election year. These types of models are designed more to help us obtain an understanding of the correlation between fundamental data and election outcomes, than for forecasting the election during the cycle. Third, most fundamental data models forecast just the presidential national popular vote; examples include those of Abramowitz (2004, 2008). This is a serious issue, not just because it is not the ideal outcome variable, but because it means that there is an extremely limited identification in just one outcome every four years. Fourth, Klarner (2008) pushed the literature forward into the realm of earlier state-by-state forecasts, but still incorporated early polling in the model. In order to compare the value of the different data sources, it is crucial to consider models that use only one data source. Without any polling data, improving on the variable choice and range of data, the model presented by Hummel and Rothschild (2013) has much smaller errors than that of Klarner (2008), and could be put to use by June 15 of the election year. Thus, I utilize the model of Hummel and Rothschild (2013) exclusively as the fundamental model for this paper, because it is the most accurate state-by-state fundamental model for Electoral College and senatorial elections, can be executed early in the cycle, and excludes voter intention polling data. The out-of-sample errors for the model of Hummel and Rothschild (2013) are smaller than the within-sample errors for the most widely circulated state-by-state fundamental models, including Klarner's most recently updated model (Klarner, 2012).[1]

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

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

[1] Klarner (2012) drops the use of voter intention polling data, which were used in early versions of the model; however, his paper was not released until after the initial running of the model for and circulation of this paper, as I wanted to ensure that 2012 would be completely out-of-sample; that change brought Klarner's model closer to that of Hummel and Rothschild (2013), which was already available in a widely circulated working paper at that time. Still, while Hummel and Rothschild (2013) had similar errors for the estimated vote share to those of Klarner (2012) for the Electoral College in 2012, it had significantly smaller errors, nearly a full point on average, for the senatorial elections.