MIT MANAGEMENT SLOAN SCHOOL

MIT Sloan School of Management

MIT Sloan School Working Paper 6166-20

The Past as Prologue: A New Approach to Forecasting

Megan Czasonis, Mark Kritzman, and David Turkington

January 20, 2021

THE PAST AS PROLOGUE:

A NEW APPROACH TO FORECASTING

Megan Czasonis
mczasonis@statestreet.com

Mark Kritzman
kritzman@mit.edu

David Turkington
dturkington@statestreet.com

FIRST VERSION: AUGUST 10, 2020

THIS VERSION: JANUARY 20, 2021

Abstract

It is common practice to forecast social, political, and economic outcomes by polling people about their intentions. This approach is direct, but it can be unreliable in settings where it is hard to identify a representative sample, or where subjects have an incentive to conceal their true intentions or beliefs. The authors propose that, as a substitute or a supplement, forecasters use historical outcomes to predict future ones. The relevance of historical events, however, is not guaranteed. The authors apply a novel technique called Partial Sample Regression to identify, in a mathematically precise way, the subset of events that are most relevant to the present. The outcomes of those events are then weighted by their relevance and averaged to give a prediction for the future. The authors illustrate their technique by showing that it correctly predicted the winner of the last six U.S. presidential elections based only on the political, geopolitical, and economic circumstances of the election year.

THE PAST AS PROLOGUE:
A NEW APPROACH TO FORECASTING

Introduction

In the late 1980's Princeton economist Orley Ashenfelter put forth a simple regression equation to predict the quality of wine 10 years in the future.[1] He used the amount of winter rainfall and rainfall during harvest, along with the average temperature during the growing season, as predictors of the future quality of a vintage, which he measured as a wine's price at auction. Notably, he did not include any variables about winemaking decisions, such as when to harvest the grapes, how to blend the varietals, or how long to keep the wine in barrels. Nor did he include information about ratings from wine critics, such as Robert Parker. Although many observers were amused by Ashenfelter's implicit challenge to winemakers and wine critics, his experiment revealed a fundamental, yet overlooked, feature of forecasting – that there are often pre-existing conditions or circumstances beyond anyone's control that strongly influence outcomes.

In the spirit of Ashenfelter, we argue that forecasts of social, political, and economic outcomes should consider how similar circumstances have affected outcomes in the past. This approach offers a supplement to public opinion polls and survey-based research. Polling is the most obvious way to gather data on what people will do, but it is not always reliable. For example, polls are widely regarded as failing to deliver accurate predictions of the 2016 and 2020 U.S. presidential elections. In its publicly available report following the 2016 election,[2] the American Association for Public Opinion Research (AAPOR) – a professional organization

comprising researchers and practitioners from academia, media, government, and the private sector – highlights multiple issues that may have led to 2016's polling bias. First off, the sampling may have misrepresented voter turnout and voter beliefs by over-weighting younger and more highly educated citizens who are more likely to answer surveys. Another possibility is the "shy Trump" theory that some voters misreported or otherwise concealed their support for the Republican nominee. In addition, some voters may have waited until the last minute to decide, in which case their full consideration of circumstances would not have been apparent to pollsters. That the same issues are likely to blame for errors in 2020 highlights how difficult it is to adjust for bias even when it is widely acknowledged.

The key advantage of polls and surveys is that they directly relate to the subject of the forecast. The same is not true for historical events – we must gauge each observation's relevance in light of current circumstances. Unlike Ashenfelter who used a simple forecasting technique to capture the impact of initial conditions, which for his purposes was sufficient, we assert that initial conditions often matter in a complex way that we can only grasp with more sophisticated quantitative tools. We therefore apply a novel forecasting technique called Partial Sample Regression.

The essence of Partial Sample Regression is best explained by first considering the implicit assumption of linear regression analysis. A linear regression equation applied to time series data is fitted based on the notion that whatever occurred during relevant periods in history will recur, and whatever occurred during non-relevant periods in history will recur but in the opposite direction. Relevance, within this context, has a precise mathematical meaning, which we will discuss shortly. This implicit assumption of linear regression analysis invites a

fundamental question. Are relevant and non-relevant observations equally useful? Linear regression analysis assumes they are. Partial Sample Regression is based on the premise that one can often derive a more reliable forecast by focusing more on relevant observations and less on non-relevant observations.

In some cases, the implicit assumption of linear regression analysis makes sense. This might be true of Ashenfelter's wine forecasting model. If the current season had above average temperatures and rainfall, past seasons with below average temperatures and rainfall might be just as relevant to forecasting the quality of this season's wine as past seasons with above average temperatures and rainfall. Suppose instead we wish to forecast the likelihood of an economic recession. Most economists would agree that the circumstances of past recessions are more relevant to forecasting the likelihood of a future recession than the circumstances of past periods of robust growth.

The forecasting methodology we propose can be applied broadly throughout the social sciences. It is suitable for any forecasting situation in which relevant observations might be more useful to a forecast than non-relevant observations, again keeping in mind that relevance has a special meaning. In this paper, we offer a detailed illustration of our approach in the context of predicting U.S. presidential elections. We show that it would have correctly predicted the winner of all five elections from 2000 and 2016 using information available at the time, and that its out-of-sample prediction for the 2020 election – which we included in this paper's first version before the election took place – also turned out to be correct. Our application to elections also highlights the intuitive appeal of extrapolating from past events. We encourage practitioners to combine our data-driven historical approach with their own

subjective views of historical relevance and with data that is available from polls, surveys, or other sources.

Partial Sample Regression

Partial Sample Regression is a two-step process. We first fit a linear regression equation to identify a subset of relevant historical observations. We then produce our forecast by invoking an obscure mathematical equivalence – that the prediction from a linear regression equation is equal to the weighted average of the past values of the dependent variable in which the weights are the relevance of the past observations for the independent variables.[3]

Relevance

As we alluded to earlier, relevance has a precise meaning. It is the sum of the statistical similarity of a particular historical observation x_i to the current values x_t for the independent variables, which is the negative of the Mahalanobis distance between them, and the informativeness of that historical observation x_i, which equals its Mahalanobis distance from the average values $\bar{x}$ of the independent variables.

Equation 1 measures the multivariate similarity between x_i and x_t.

$$similarity(x_i, x_t) = -(x_i - x_t)\Omega^{-1}(x_i - x_t)' \quad (1)$$

Here x_t is a row vector of the current observations of the independent variables, x_i is a row vector of the prior observations of the independent variables, the symbol $'$ indicates matrix transpose, and Ω^{-1} is the inverse covariance matrix of X where X comprises all the vectors of the independent variables. In contrast to other distance measures, the Mahalanobis distance measure considers not only how independently similar the prior observations of the independent variables are to the current observations but also the similarity of their interaction to the interaction of the current observations. In other words, the Mahalanobis distance judges the similarity of two multivariate observations based on the patterns of co-occurrence of the variables rather than their typical patterns of co-occurrence across the entire historical sample. All else equal, prior observations for the independent variables that are more like the current observations are more relevant than prior observations that are less similar. Equation 1 can be thought of as a mathematical formalization of the common practice of looking for past experiences that are like current conditions to form a prediction.

However, not all similar observations are alike. Observations that are close to their historical averages may be driven more by noise than by events. These ordinary occurrences are therefore less relevant. Observations that are distant from their historical averages are unusual and therefore more likely to be driven by significant events. These event-driven observations are potentially more informative.[4] Given this intuition, we define the informativeness of a prior observation x_i as its Mahalanobis distance from its average value, $\bar{x}$.

$$informativeness(x_i) = (x_i - \bar{x})\Omega^{-1}(x_i - \bar{x})' \quad (2)$$

The relevance of an observation x_i is equal to the sum of its multivariate similarity and its informativeness.

$$relevance(x_i; x_t) = similarity(x_i, x_t) + informativeness(x_i) \quad (3)$$

To summarize, similarity equals the negative of the Mahalanobis distance of a prior observation of x_i from the current observation x_t. Informativeness equals the Mahalanobis distance of x_i from its historical average. Relevance equals the sum of similarity and informativeness. In other words, prior periods that are like the current period but are different from the historical average are more relevant than those that are not.

Equivalence of $\hat{y}_t$ to Relevance-Weighted Average of Past y_is

Our definition of relevance is hardly arbitrary. It turns out that it is mathematically equivalent to interpret $\hat{y}_t$ from a fitted Ordinary Least Squares (OLS) linear regression equation as the weighted average of the prior y_is in which the weights equal the relevance of the prior x_is. To see this equivalence, we restate Equation 3 using the notation of Equations 1 and 2 and then simplify, as shown in Equations 4 and 5.[5]

$$relevance\ (x_i; x_t) = -(x_i - x_t)\Omega^{-1}(x_i - x_t)' + x_i\Omega^{-1}x_i' \quad (4)$$

$$relevance(x_i; x_t) = 2x_t\Omega^{-1}x_i' - x_t\Omega^{-1}x_t' \quad (5)$$

We then express the prediction $\hat{y}_t$ from a fitted regression equation as a relevance-weighted average of prior observations for y_i times a simple scalar multiple of 1/2:

$$\hat{y}_t = \frac{1}{2} \cdot \frac{1}{N} \sum_{i=1}^{N} relevance\ (x_i; x_t) y_i \tag{6}$$

$$\hat{y}_t = \frac{1}{2N} \sum_{i=1}^{N} 2 x_t \Omega^{-1} x_i' y_i - x_t \Omega^{-1} x_t' y_i \tag{7}$$

For simplicity in this illustration, we shift the observed y_is to have a mean value of zero, which causes the final term to drop out yielding:

$$\hat{y}_t = \frac{1}{N} \sum_{i=1}^{N} x_t \Omega^{-1} x_i' y_i \tag{8}$$

The expression for the covariance matrix (remembering that we have assumed means of zero for each component of X) is given by:

$$\Omega = \frac{1}{N} X'X \tag{9}$$

The inverse of the covariance matrix is given by:

$$\Omega^{-1} = N(X'X)^{-1} \tag{10}$$

Substituting this value into the expression for $\hat{y}_t$, we have:

$$\hat{y}_t = \sum_{i=1}^{N} x_t (X'X)^{-1} x_i' y_i \tag{11}$$

By expressing Equation 11 in standard matrix notation, we have:

$$\hat{y}_t = x_t (X'X)^{-1} X'Y \tag{12}$$

This leads us to the familiar standard solution for generating a prediction from a fitted linear regression:

$$\hat{y}_t = x_t \beta' \tag{13}$$

Once we determine the relevance of the past observations of the independent variables, we use the equivalence that we just demonstrated to form a prediction from a subsample of the most relevant historical observations. We compute our prediction from the observations in this subsample as a weighted average of the past values of the dependent variable in which the weights are the relevance of the past observations for the independent variables.

Equation 14 gives our prediction equation, where $\bar{y}$ is the equally weighted average of the n data points in the relevant subsample. If we were to include every available historical observation, this formula would yield precisely the same prediction as a typical linear regression model.

$$\hat{y}_t(x_t) = \bar{y} + \frac{1}{2n}\sum_{i=1}^{n} relevance(x_i; x_t)(y_i - \bar{y}) \tag{14}$$

Data and Methodology

Our historical sample includes 35 presidential elections from 1876 through 2020. For each election, we collect a range of political, geopolitical, and economic data as our predictors. We pool this historical election data and apply Partial Sample Regression to predict the outcome of the last six presidential elections (2000, 2004, 2008, 2012, 2016 and 2020). For each election, we base our prediction on the 50% most relevant prior elections, where relevance is defined as

the sum of multivariate similarity and informativeness. All our predictions are out-of-sample, based only on data available as of July 31st of the election year, and accounting for the point-in-time economic data that would have been available prior to subsequent revisions. Our dependent and independent variables are described below.

Dependent variable

- Percentage of electoral votes for the Democratic candidate.
  Source: "United States Presidential Election Results," Encyclopaedia Britannica.

Prior to running the regression, we must perform a two-step transformation of the dependent variable to prevent our model from predicting an outcome that violates the bounds of zero and one for the percentage of electoral votes won by the Democratic candidate. We first use the logit transformation to convert the percentage of electoral votes to range from negative infinity to positive infinity. This transformation is as follows:

$$y = logit(Y) = log\left(\frac{Y}{1-Y}\right) \quad (15)$$

In Equation 15, lower case y is the logit of the percentage of electoral votes, which ranges from negative infinity to positive infinity, and upper case Y is the actual percentage of electoral votes, which ranges from zero to one. After running the regression, we transform the resulting prediction back so that it is a percentage ranging from zero to one:

$$\hat{Y} = \frac{1}{1+ \exp(-\hat{y})} \quad (16)$$

Here $\hat{Y}$ is the predicted percentage of electoral votes, which ranges from zero to one, and $\hat{y}$ is its logit, which ranges from negative infinity to positive infinity.

Independent variables

Our independent variables fall into three categories: political, geopolitical, and economic variables.

Political variables:

- Party affiliation of incumbent president (0 or 1)
  Source: "By Political Party," Potus.com
- Is the incumbent running for another term? (0 or 1)
- Senate – Majority party (0 or 1)
  Source: "Party Division," Senate.gov
- Senate – Percentage of seats held by Democrats
  Source: "Party Division," Senate.gov
- House – Majority party (0 or 1)
  Source: "Party Division of the House of Representatives, 1789 to Present," History.house.gov
- House – Percentage of seats held by Democrats
  Source: "Party Division of the House of Representatives, 1789 to Present," History.house.gov

Geopolitical variable:

- Was the U.S. at war during the election year? (0 or 1)
  Source: "America's Wars," Department of Veterans Affairs and "U.S. Periods of War and Dates of Recent Conflicts," Congressional Research Service, (0 or 1)

Economic variables:

- Was the U.S. in a recession during the election year? (0 or 1)
  Source: NBER.org
- Trailing four-year economic growth, measured as percentage change in GDP
  Source: Jorda-Schularick-Taylor Macrohistory Database, microhistory.net
- Trailing four-year change in debt, measured as change in Debt-to-GDP
  Source: Jorda-Schularick-Taylor Macrohistory Database, microhistory.net
- Trailing four-year US stock return
  Source: Jorda-Schularick-Taylor Macrohistory Database, microhistory.net

Confidence Bands

When we forecast an election, we would like to know not only the weighted average forecast derived from the most relevant observations, but also how much agreement there is across the observations that comprise the weighted average. Specifically, we want to know to what extent the most relevant observations for past election outcomes are like each other. Holding all else equal, we should have more confidence in a forecast that arises from a collection of similar election outcomes than in the same forecast that arises from a very dispersed range of outcomes. We obtain this measure by estimating the standard deviation of the prediction based on the subsample that feeds it. The result of this calculation, which we describe in

Appendix B, is a standard error akin to the margin of error that is typically reported along with polling data.

Results

Exhibit 1 reports our model's predictions and actual outcomes for the percentage of electoral votes won by the Democratic candidate for each election since 2000. For comparison, we also report polling results as of November of the election year. Our model correctly predicted the presidential victor in all six elections since 2000. Notably, it correctly predicted the 2016 outcome, which polls failed to anticipate. Moreover, in most cases, the model gives a clear projection, more so than polls which tend to oscillate around 50%. For example, in 2004, our model predicted a clear Republican victory while public polling was undecided at 50%.

Exhibit 1: Out-of-Sample Predictions and Realizations
(Percentage of electoral votes for Democratic candidate)

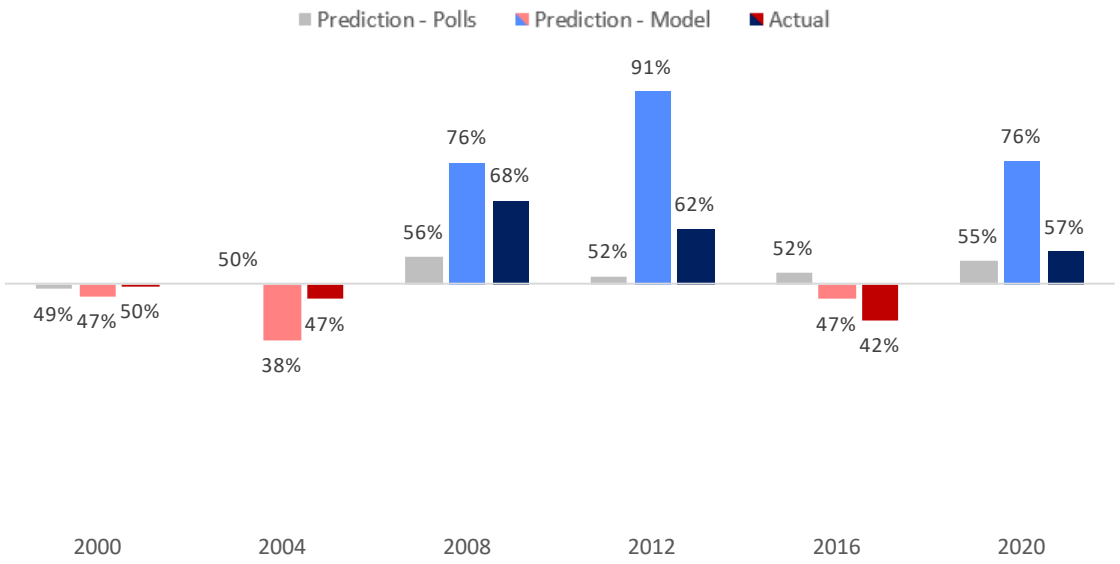

Exhibit 2 shows the one standard deviation confidence bands of our model's predictions. These confidence bands are analogous to the margin of error that is commonly reported along with polling results.

Exhibit 2 shows that our model predicted marginal Republican victories in 2000 and 2016, though it was more confident in its 2016 prediction. Moreover, our model was relatively confident in the direction of the result for 2008, 2012, and 2020.

Exhibit 2: Confidence Bands of Model Predictions

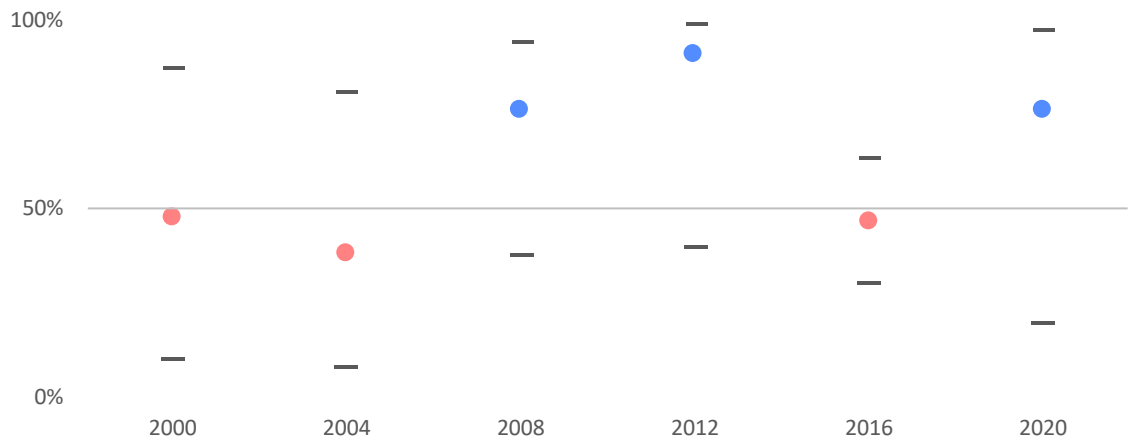

Exhibit 3 provides a detailed view of the relevant subsample of elections that are used in each prediction. The height of each circle equals the percentage of electoral votes for the

Democratic candidate (with color indicating the winner), and the area of each circle is proportional to the relevance of that observation.

Exhibit 3: Statistically Relevant Prior Elections and their Outcomes

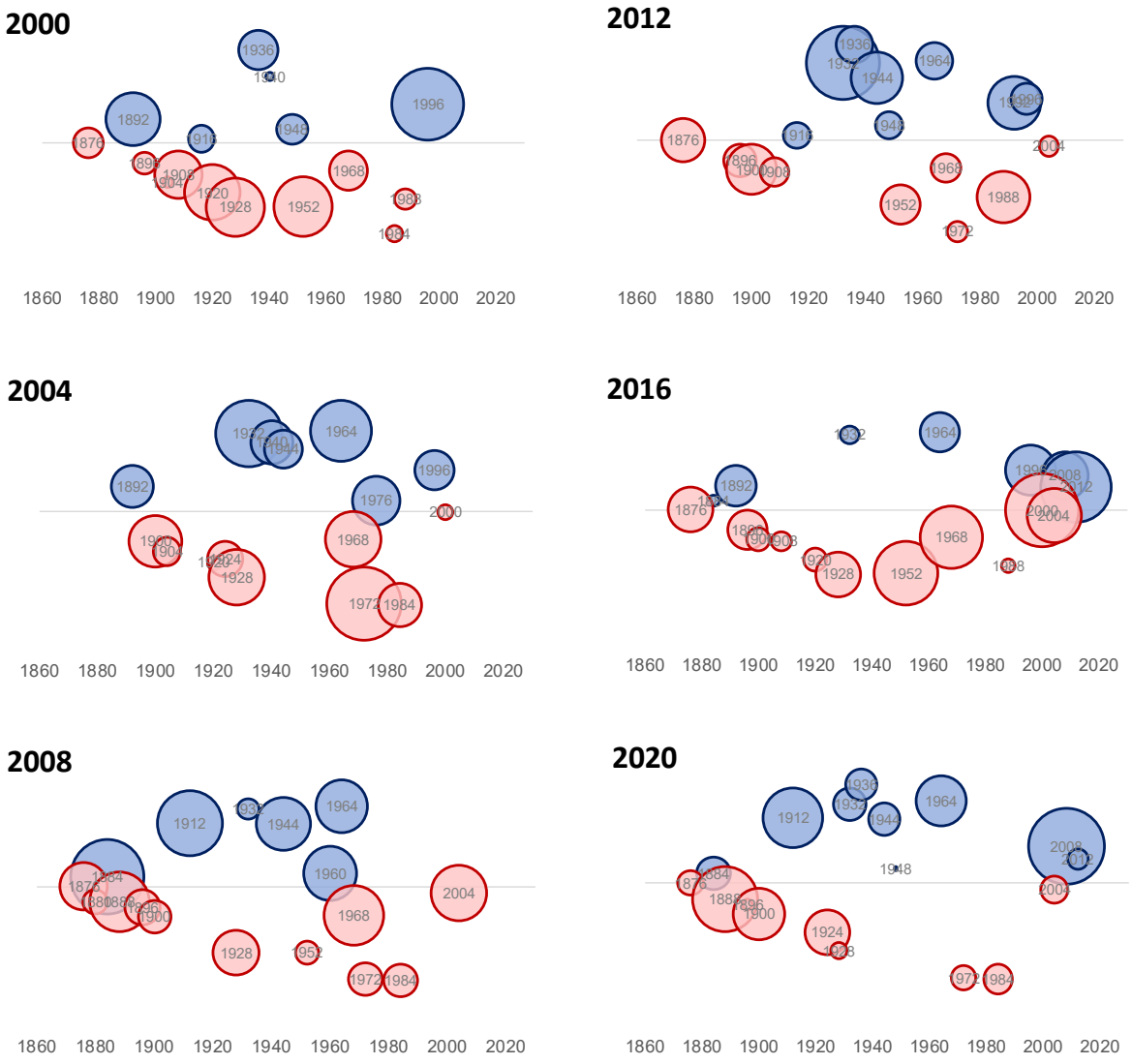

Exhibit 3 reveals some interesting observations. For the 2000 election, which the model correctly predicted as a Republican victory, 1996 was the most statistically relevant election, though it had a Democratic winner. For the 2020 election, 2008 is the most relevant past

election. It is worth noting that the 2008 election occurred at the inception of the Global Financial Crisis and the 2020 election also occurred during an economic and financial crisis.

In 2004, the results of similar past elections varied widely. The largest weights correspond to blowout victories for both parties, most notably including Nixon's reelection in 1972. This dramatic variation caused the wide error bands we observed earlier for the 2004 prediction. Interestingly, George Bush's win over Al Gore just four years prior was not considered very relevant, perhaps due to the 9/11 terrorist attacks and the dotcom stock market crash that occurred in the interim. By contrast, the 2016 prediction judged that same 2000 election to be the most relevant, followed by 2012, which paints a mixed picture. However, when we look back further in history, most of the other relevant elections were Republican victories, which tilted the model to predict a Republican win with a tight distribution. The most relevant observation for the 2020 prediction was Barack Obama's first win in 2008. That year shared characteristics with 2020 including significant negative economic shocks and a similar alignment of political power in Congress, the House of Representatives, and the White House. The next most relevant observations occurred more than 100 years in the past.

Summary

We invoke two under recognized principles of forecasting to predict the outcomes of presidential elections, and we present a mathematical formalization of the second principle. The first principle is that outcomes often depend on pre-existing conditions that are beyond

anyone's control, and that these conditions should be considered in a forecasting model. The second principle, which contradicts the basic premise of linear regression analysis, is that it is often better to base a prediction on a subset of more relevant observations than on the full sample of observations.

Forecasters often abide by this second principle, but they typically do so informally. For example, when political scientists or pundits forecast presidential elections, they often analyze past elections for clues about upcoming elections. But they do not treat all past elections alike. They judge some to be more relevant than others. This behavior is true in general when we try to predict an outcome based on prior experiences. We look for those events that bear some resemblance to current conditions. We apply this concept in a mathematically formal way, which requires us to incorporate a second, and perhaps less obvious, component of relevance. We consider the unusualness of the past experiences. The intuition is that unusual occurrences are more informative than common occurrences, which simply might be a manifestation of noise in the data.

Once we identify a subsample of relevant historical elections, we invoke an obscure mathematical equivalence to form our predictions. The prediction from a linear regression equation equals a weighted average of the past values of the dependent variable in which the weights are the relevance of the values for the independent variables. We apply this equivalence to our relevant subsample of political, geopolitical, and economic data to form our predictions.

We show out-of-sample predictions and realizations of the past six U.S. presidential elections along with polling data and confidence bands. For each prediction, we also attribute the weight our model places on each prior election it includes. Our model correctly predicted the outcomes of the past six elections.

Our forecast of presidential elections is but one illustration of the methodology described in this paper. This methodology could be applied broadly throughout the social sciences, especially for situations in which relevant observations might be more useful to a forecast than non-relevant observations.

Appendix A: Model Inputs

Inputs (page 1)

	Candidates		Outcome		Incumbent		
Election	D	R	Winner	Electoral votes to D		Party	Running?
1876	Samuel J. Tilden	Rutherford B. Hayes	R	50%	Ulysses S. Grant	R	N
1880	Winfield Scott Hancock	James A. Garfield	R	42%	Rutherford B. Hayes	R	N
1884	Grover Cleveland	James G. Blaine	D	55%	Chester A. Arthur	R	N
1888	Grover Cleveland	Benjamin Harrison	R	42%	Grover Cleveland	D	Y
1892	Grover Cleveland	Benjamin Harrison	D	62%	Benjamin Harrison	R	Y
1896	William Jennings Bryan	William McKinley	R	39%	Grover Cleveland	D	N
1900	William Jennings Bryan	William McKinley	R	35%	William McKinley	R	Y
1904	Alton B. Parker	Theodore Roosevelt	R	29%	Theodore Roosevelt	R	Y
1908	William Jennings Bryan	William Howard Taft	R	34%	Theodore Roosevelt	R	N
1912	Woodrow Wilson	William Howard Taft	D	82%	William Howard Taft	R	Y
1916	Woodrow Wilson	Charles Evans Hughes	D	52%	Woodrow Wilson	D	Y
1920	James M. Cox	Warren G. Harding	R	24%	Woodrow Wilson	D	N
1924	John W. Davis	Calvin Coolidge	R	26%	Calvin Coolidge	R	Y
1928	Alfred E. Smith	Herbert Hoover	R	16%	Calvin Coolidge	R	N
1932	Franklin D. Roosevelt	Herbert Hoover	D	89%	Herbert Hoover	R	Y
1936	Franklin D. Roosevelt	Alfred M. Landon	D	98%	Franklin D. Roosevelt	D	Y
1940	Franklin D. Roosevelt	Wendell L. Willkie	D	85%	Franklin D. Roosevelt	D	Y
1944	Franklin D. Roosevelt	Thomas E. Dewey	D	81%	Franklin D. Roosevelt	D	Y
1948	Harry S. Truman	Thomas E. Dewey	D	57%	Harry S. Truman	D	Y
1952	Adlai E. Stevenson	Dwight D. Eisenhower	R	17%	Harry S. Truman	D	N
1956	Adlai E. Stevenson	Dwight D. Eisenhower	R	14%	Dwight D. Eisenhower	R	Y
1960	John F. Kennedy	Richard M. Nixon	D	56%	Dwight D. Eisenhower	R	N
1964	Lyndon B. Johnson	Barry M. Goldwater	D	90%	Lyndon B. Johnson	D	Y
1968	Hubert H. Humphrey	Richard M. Nixon	R	36%	Lyndon B. Johnson	D	N
1972	George S. McGovern	Richard M. Nixon	R	3%	Richard M. Nixon	R	Y
1976	Jimmy Carter	Gerald R. Ford	D	55%	Gerald R. Ford	R	Y
1980	Jimmy Carter	Ronald W. Reagan	R	9%	Jimmy Carter	D	Y
1984	Walter F. Mondale	Ronald W. Reagan	R	2%	Ronald W. Reagan	R	Y
1988	Michael S. Dukakis	George H.W. Bush	R	21%	Ronald W. Reagan	R	N
1992	Bill Clinton	George H.W. Bush	D	69%	George H.W. Bush	R	Y
1996	Bill Clinton	Bob Dole	D	70%	Bill Clinton	D	Y
2000	Al Gore	George W. Bush	R	50%	Bill Clinton	D	N
2004	John Kerry	George W. Bush	R	47%	George W. Bush	R	Y
2008	Barack Obama	John McCain	D	68%	George W. Bush	R	N
2012	Barack Obama	Mitt Romney	D	62%	Barack Obama	D	Y
2016	Hillary Clinton	Donald Trump	R	42%	Barack Obama	D	N
2020	Joe Biden	Donald Trump	D	57%	Donald Trump	R	Y

* 2020 economic and market variables based on data available as of July 31, 2020.

Inputs (page 2)

	Congress				War	Economy and stock market			
	Senate		House				4-yr	4-yr change	4-yr
Election	Majority	% D	Majority	% D	At war?	Recession?	GDP growth	in debt-to-GDP	stock return
1876	R	37%	D	62%	N	Y	1.0%	-0.01	-5.7%
1880	D	55%	D	48%	N	N	24.7%	-0.05	97.5%
1884	R	47%	D	60%	N	Y	13.7%	-0.06	-6.1%
1888	R	49%	D	51%	N	Y	17.7%	-0.04	40.9%
1892	R	44%	D	72%	N	N	18.0%	-0.04	27.0%
1896	R	44%	R	26%	N	Y	-5.3%	0.02	-7.6%
1900	R	29%	R	45%	Y	Y	32.7%	-0.02	89.0%
1904	R	37%	R	46%	N	Y	24.9%	-0.02	42.1%
1908	R	34%	R	43%	N	Y	17.3%	-0.01	31.3%
1912	R	46%	D	58%	N	Y	24.1%	-0.01	25.9%
1916	D	58%	D	53%	N	N	32.8%	-0.01	29.3%
1920	R	49%	R	44%	N	Y	78.1%	0.25	-6.5%
1924	R	44%	R	48%	N	Y	-1.6%	-0.03	87.8%
1928	R	48%	R	45%	N	N	12.0%	-0.06	171.7%
1932	R	49%	R	50%	N	Y	-39.5%	0.15	-61.4%
1936	D	72%	D	74%	N	N	42.7%	0.07	195.7%
1940	D	72%	D	60%	N	N	21.2%	0.02	-23.1%
1944	D	59%	D	51%	Y	N	118.3%	0.48	57.8%
1948	R	47%	R	43%	N	Y	22.4%	0.07	41.2%
1952	D	51%	D	54%	Y	N	33.8%	-0.24	120.0%
1956	D	50%	D	53%	N	N	22.4%	-0.10	111.8%
1960	D	65%	D	65%	N	Y	20.7%	-0.08	40.7%
1964	D	66%	D	59%	Y	N	26.2%	-0.07	66.8%
1968	D	64%	D	57%	Y	N	37.4%	-0.07	43.4%
1972	D	54%	D	59%	Y	N	36.1%	-0.05	25.1%
1976	D	61%	D	67%	N	N	46.4%	-0.01	5.2%
1980	D	58%	D	64%	N	Y	52.5%	-0.03	55.1%
1984	R	45%	D	62%	N	N	41.2%	0.07	48.6%
1988	D	55%	D	59%	N	N	30.0%	0.11	93.5%
1992	D	56%	D	61%	N	N	24.5%	0.12	78.8%
1996	R	48%	R	47%	N	N	23.9%	0.03	88.0%
2000	R	45%	R	49%	N	N	27.0%	-0.09	88.9%
2004	R	48%	R	47%	Y	N	19.3%	0.05	-4.0%
2008	R	49%	D	54%	Y	Y	19.9%	0.07	-20.3%
2012	D	51%	R	44%	Y	N	9.8%	0.32	75.7%
2016	R	44%	R	43%	Y	N	15.3%	0.06	53.2%
2020*	R	45%	D	54%	Y	Y	2.0%	0.03	48.4%

* 2020 economic and market variables based on data available as of July 31, 2020.

Appendix B: Computing standard errors

We begin with the formula for Partial Sample Regression, as restated below, in which we sum over any chosen number n of the most relevant observations. Recall that the forecast $\hat{y}_t$ is a relevance-weighted sum of historical observations y_i in which the relevance of each historical observation is defined as in Equation 4 and notated here as r_{it} with respect to the current conditions, x_t. For completeness, we also put $n-1$ in the denominator, which yields the unbiased forecast as compared to simply dividing by n. In practice, this distinction is not material.

$$\hat{y}_t(x_t) = \bar{y} + \frac{1}{2(n-1)} \sum_{i=1}^{n} r_{it}(y_i - \bar{y}) \tag{17}$$

Next, we rewrite the prediction formula equivalently as a weighted sum of y_i:

$$\hat{y}_t(x_t) = \frac{1}{2(n-1)} \sum_{i=1}^{n} \left(\frac{2(n-1)}{n} + r_{it} - \bar{r} \right) y_i \tag{18}$$

We express and simplify the standard deviation of a given prediction, $\hat{y}_t$, as follows:

$$\sigma_{\hat{y}_t}^2 = E[(\hat{y}_t - \bar{y})^2] \tag{19}$$

$$\sigma_{\hat{y}_t}^2 = E\left[\left(\frac{1}{2(n-1)}\sum_{i=1}^{n}\left(\frac{2(n-1)}{n} + r_{it} - \bar{r}\right)y_i\right)^2\right] - \bar{y}^2 \quad (20)$$

$$\sigma_{\hat{y}_t}^2 = \frac{1}{4(n-1)^2}\sum_{i=1}^{n}\left(\frac{2(n-1)}{n} + r_{it} - \bar{r}\right)^2 E[y_i^2] - \bar{y}^2 \quad (21)$$

$$\sigma_{\hat{y}_t}^2 = \frac{\sigma_{\hat{y}}^2}{n} + \frac{n}{4(n-1)^2}\sigma_r^2(\sigma_y^2 + \bar{y}^2) \quad (22)$$

Note that Equation 21 follows from Equation 22 because we assume that $E[\frac{(y_i-\bar{y})}{\sigma_y}\frac{(y_j-\bar{y})}{\sigma_y}] = 0$ when i does not equal j, as there should not be a systematic relationship in the deviations of y across every pair of observations that occur at different times. In practice, even if we estimate these cross-observation relationships empirically, the term is nearly zero and it is immaterial. If the variance of relevance scores, σ_r^2, were zero, our prediction would be the mean of the subsample and its standard error would equal the familiar expression for the standard error of the mean, $\sigma_{\hat{y}_t} = \frac{\sigma_y}{\sqrt{n}}$. All else equal, the standard error of our prediction is lower when there are more data points, higher when there is more variation in the dependent variable, and higher when there is more variation in relevance within the partial sample.

This analysis leads to the insight that prediction confidence depends on the conditions x_t that prevail at the time of the prediction. The same model using the same data might be capable of rendering a confident prediction in one circumstance, but not in another. Prediction confidence depends on the availability of relevant historical data as well as the consistency of the relationship that occurred in the relevant sample.

Notes

This material is for informational purposes only. The views expressed in this material are the views of the authors, are provided "as-is" at the time of first publication, are not intended for distribution to any person or entity in any jurisdiction where such distribution or use would be contrary to applicable law and are not an offer or solicitation to buy or sell securities or any product. The views expressed do not necessarily represent the views of Windham Capital Management, State Street Global Markets®, or State Street Corporation® and its affiliates.

References

"America's Wars," *Department of Veterans Affairs,* 2019. U.S. Department of Veterans Affairs. Web. 29 July 2020.

"By Political Party." *Potus.com*. Web. 29 July 2020.

Chow, G., E. Jacquier, M. Kritzman, and K. Lowry. 1999. "Optimal Portfolios in Good Times and Bad." *Financial Analysts Journal*, vol. 55, no. 3 65-71 (May/June).

Czasonis, M., M. Kritzman and D. Turkington. 2020. "Addition by Subtraction: A Better Way to Predict Factor Returns – And Everything Else." *The Journal of Portfolio Management*, vol. 46, no. 8 (September).

Czasonis, M., M. Kritzman, and D. Turkington. 2021. "The Stock-Bond Correlation." *Forthcoming in the Journal of Portfolio Management* (February).

Òscar Jordà, Moritz Schularick, and Alan M. Taylor. 2017. "Macrofinancial History and the New Business Cycle Facts." in *NBER Macroeconomics Annual 2016*, volume 31, edited by Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press.

"Party Division." *Senate.gov*. United States Senate. Web. 29 July 2020.

"Party Division of the House of Representatives, 1789 to Present." *History.house.gov*. History, Art & Archives, U.S. House of Representatives. Web. 29 July 2020.

Peter Passell. 1990. "Wine Equation Puts Some Noses Out of Joint," New York Times, Section 1, page 1. March 4, 1990.

"United States Presidential Election Results." *Encyclopaedia Britannica*. Encyclopaedia Britannica, Inc., 2017. *Encyclopaedia Britannica*. Web. 29 July 2020.

"U.S. Periods of War and Dates of Recent Conflicts." *Congressional Research Service,* 2020. Federation of American Scientists. Web. 29 July 2020.

[1] Ashenfelter's equation and the controversy it provoked is described in a New York Times article from March 4, 1990 (Section 1, page 1) called "Wine Equation Puts Some Noses Out of Joint," by Peter Passell.
[2] https://www.aapor.org/Education-Resources/Reports/An-Evaluation-of-2016-Election-Polls-in-the-U-S.aspx
[3] See Czasonis, Kritzman, and Turkington (2020) for a detailed description of this methodology.
[4] For further discussion of noise-driven versus event-driven observations and their relationship to estimating risk, see Chow, G., E. Jacquier, M. Kritzman, and K. Lowry (1999).
[5] For notational simplicity, we assume here, without loss of generality, that the means of the independent variables equal zero.