REPORT

POLITICAL SCIENCE

Improving election prediction internationally

Ryan Kennedy,[1]* Stefan Wojcik,[2,3] David Lazer[2,3]

This study reports the results of a multiyear program to predict direct executive elections in a variety of countries from globally pooled data. We developed prediction models by means of an election data set covering 86 countries and more than 500 elections, and a separate data set with extensive polling data from 146 election rounds. We also participated in two live forecasting experiments. Our models correctly predicted 80 to 90% of elections in out-of-sample tests. The results suggest that global elections can be successfully modeled and that they are likely to become more predictable as more information becomes available in future elections. The results provide strong evidence for the impact of political institutions and incumbent advantage. They also provide evidence to support contentions about the importance of international linkage and aid. Direct evidence for economic indicators as predictors of election outcomes is relatively weak. The results suggest that, with some adjustments, global polling is a robust predictor of election outcomes, even in developing states. Implications of these findings after the latest U.S. presidential election are discussed.

Forecasting of U.S. presidential elections has been a cottage industry among political scientists and pundits for decades (*1–5*) and has also been attempted in some areas of Western Europe, including Italy (*6*, *7*), France (*8*, *9*), Spain (*10*), Great Britain (*11*, *12*), the Netherlands (*13*), and Germany (*14*, *15*). This tradition of predicting elections, primarily in long-standing democracies, raises the question of the extent to which global models for election prediction can be produced, and the extent to which election predictions can be improved, by pooling information from across a wide variety of political systems. Scholars in this area have reached very different conclusions about the feasibility of such a project (*16*, *17*).

Academic analyses of elections can be classified into four categories: structuralists, aggregators, synthesizers, and judges (*18*). Structuralists use macro features such as economic indicators to calculate the expected support for the incumbent party candidate and/or their expected vote share. These analysts use long-term, slow-moving measures with relatively low variability and high levels of comparability from election to election (*18–21*). Aggregators use national opinion poll data much more heavily than their structural counterparts, and sometimes use aggregated information from election markets as well (*22*). Synthesizers combine the tools of aggregators and structuralists to make their forecasts (*23*). Finally, judges use various quantitative and nonquantitative techniques to evaluate the existing evidence through their own expert lens (*24*). However, the major criticism of contemporary forecasting methods in the academy is that the data are of poor quality, and the number of observations is so small that the signal-to-noise ratio is too low to make good forecasts (*25*).

After successfully predicting the 2012 election results in all 50 states, Nate Silver was crowned the king of election forecasting in the United States, and other approaches quietly did just as well. Drew Linzer's Votamatic and Simon Jackman's uniform swing model performed roughly as well as Nate Silver's model in the 2012 elections (*26*). These models generally fall into the synthesizer category, using a combination of aggregated polling data and structural forecasts.

The failure of these models, however, to predict Donald Trump's electoral college victory has led to a new wave of criticism of quantitative methods for predicting elections. Part of this comes from the fact that the United States, while rich in data, is poor in cases—national executive elections only come once every 4 years—and unusual in structure because of the electoral college. A forecast placing the probability of an outcome at 80% will be wrong 20% of the time, but it would take hundreds of years to validate whether the model was indeed correct 80% of the time in similar circumstances.

The issue of low numbers of cases can be addressed using cross-national data. Indeed, a smaller body of cross-national work has been aimed at analyzing numerous elections simultaneously, though such studies have generally been limited to Western Europe and North America. Moreover, they have focused more on hypothesis testing than prediction. Nadeau and colleagues analyzed 10 West European countries and revealed a strong correlation between voting, inflation, and economic perceptions (*27*). Along a similar line, Duch and Stevenson modeled voting behavior in 18 Western European countries with a "competency model," asserting that voters reward and punish politicians for prior economic conditions (*28*). The management of the economy was the best indicator of a politician's future economic stewardship, but differing political institutions obscured the ability of citizens to discern whether the politician was truly at fault for the observed economic conditions, which renders the effect of economic voting different across institutional contexts (*28*).

Global election forecasting bears numerous problems that until now have not been resolved. Most prominent among them is producing accurate global forecasts from the limited data that exist with sufficient density on a global scale (*29*). There is also the difficulty in making forecasts with sufficient lead time (*25*).

We started this project with a data set that encompassed all national-level direct executive elections in which the incumbent party could lose election from 1945 to 2012 (*30*). By direct executive election, we mean those in which voters cast ballots for the party or person who will hold executive office, as opposed to those in which the executive is indirectly elected by the ruling parliamentary coalition. The focus on direct executive elections allowed the model to concentrate on a relatively homogeneous set of institutional rules, without the extra step of parliamentary coalition formation or anticipation of minority governments, making the problem of modeling more tractable. The outcome of interest was whether the incumbent party's candidate (or the chosen successor of the current chief executive) won the subsequent election or if a candidate from an opposition party won. This outcome was reasonably comparable across a range of institutions and party systems. Cases in which the election was the first multiparty election, and therefore there was no incumbent party, were excluded. The full data for our model include 621 elections.

We divided the data into an in-sample training set that spanned from 1945 to 2006 (493 observations) and an out-of-sample test set that spanned from 2007 to 2012 (128 observations). Bayesian additive regression trees (BARTs) were used to analyze the data (*31*, *32*) (supplementary materials). BART uses the ensemble regression tree methods popular in machine learning prediction models (fig. S1), but does so within a formal probability framework, which makes it generally more stable. Our experience was that BART required far less tuning and produced more consistent results than other ensemble methods, like boosting or random forests. Tuning on the training set was done through grid-search cross-validation (supplementary materials).

Drawing from several data sets, we started with a feature set of 29 variables (table S1). Using a combination of the principled variable selection methods specific to BART and an extensive review of the literature, we settled on a feature

[1]Center for International and Comparative Studies, University of Houston, Houston, TX, USA. [2]Lazer Lab, Northeastern University, Boston, MA, USA. [3]Institute for Quantitative Social Science, Harvard University, Cambridge, MA, USA.
*Corresponding author. Email: rkennedy@uh.edu

set that included only nine variables to avoid unnecessary complexity and issues with overfitting (*32*, *33*) (figs. S2 to S4). The features used for prediction included an indicator variable for whether the current officeholder was running for the incumbent party; the level of democracy (as measured by Polity); whether a reliable poll existed before the election; whether a reliable poll predicted the incumbent party to win; whether there were concerns in the media that the election would be cancelled before the election; whether the country received a substantial amount of foreign aid; whether the country had good relations with the United States; the percentage of real gross domestic product (GDP) per-capita growth in the year before the election; and how long the current party had held the executive office. The supplementary materials provide more detail on these variables and their reason for inclusion. For observations that are missing data, we use the missing data correction recommended by (*32*).

In the training data (pre-2007), the model correctly predicted the outcome about 78.9% of the time (table S2). The model performed somewhat worse in predicting when the incumbent party is going to lose, achieving about 70.1% accuracy, which is not too surprising given what we demonstrate about the individual variables below. In the testing data (2007 and beyond), the results improve slightly, predicting 81.9% of the election outcomes correctly, again with a better record at predicting when the incumbent party is likely to win (fig. S5). By way of comparison, a naïve model, which predicts that the incumbent party always wins, would get about 60% correct in the training data and 75% correct in the test data (table. S3).

Figure 1 shows partial dependence plots of the features used in the models. These give an idea of what the direct (noninteraction) effect of the variables are. We found a strong and inverse relationship between the openness of the political regime, as indicated by Polity's democracy scores, and the probability of the incumbent party remaining in office. There was also a strong positive impact of the current officeholder running. This is consistent with the incumbent advantage literature in the U.S. context (*34*, *35*). Countries that received a large amount of economic aid were less likely to return the incumbent party to office. Given that this arises in addition to the effect of traditional economic variables, this outcome seems to support the contention that international aid assists in supporting competitive elections (*36*, *37*). Good relations with the United States increased the probability of the incumbent party holding office. Although the literature is less clear on why this would be the case, some scholars have suggested that good relations with the United States are accompanied by aid flows and political support that may help sway an election (*38*). Concerns in the media that the election would be cancelled increased the probability of the incumbent party returning to office, suggesting this as a potential prior indicator of electoral manipulation (*39*). None of the economic variables produced a strong independent effect. How long a party has controlled the executive office had a nonlinear effect, with an accumulating advantage over time. Finally, polling data were powerfully predictive, where the very existence of a reliable poll was already a (weak) indication that the incumbent might lose, and the results of these polls provided the strongest indication of the outcome.

Even with formal separation between the training and testing data set, there are always concerns about the extent to which the test set serves as a true out-of-sample test. To ensure that our models functioned outside of the lab, we participated in two live out-of-sample tests between 2013 and 2015. In these tests, we used somewhat different feature sets and, for a time, a different modeling strategy. These did not markedly alter the results above and the approach was generally comparable, but it did reflect the fact that model building and real-time data collection were being developed simultaneously with the live testing. The feature set for these models included the year of the election, whether the incumbent party candidate was the current officeholder, if the current officeholder was not running because of term limits, whether the current officeholder had artificially extended his or her term in office, the democracy score, whether a reliable poll existed, and whether the poll showed the incumbent party's candidate winning. For much of the period in which these tests were conducted, we used an adaptive boosting model and switched to the BART model in June 2014 (supplementary materials). There were also some issues with data collection, including a move to a more updated version of NELDA in June 2014 and a couple of situations in which two data-collection groups found different polling data (supplementary materials). The effect of these was negligible on our overall accuracy results, as errors were rare, occurred randomly, and were offset across the course of the study.

The first live test was a project sponsored by the Intelligence Advanced Research Projects Activity (IARPA) and involved prediction of Latin American elections in 2013 and 2014. Forecasts were submitted at least 2 weeks before the election. In this exercise, our model correctly forecast the outcome of 10 out of the 11 elections (about 90.9%) (table S6).

We also began posting live predictions for all global direct executive elections on 21 June 2013. The predictions were publicly available on the Complexity and Social Network Blog and were updated monthly for the 6 months leading to the election (supplementary materials). Data were collected by a trained team of coders, and their measurements were entered directly into the version of the algorithm active at the time (supplementary materials). The success rate for these predictions was roughly similar to that produced in the lab, getting 29 out of 36 elections correct (about 80.5%) (table S7).

Two questions naturally arise from these results. First, since we included a rough indicator for polling data, to what extent are our results dependent on having accurate polls? Although it is generally accepted that polls are likely to be more accurate than structural features, because of their ability to take into account the dynamics of the campaign, their relative accuracy compared with other forecasting tools has changed over time. For example, early betting markets performed quite well in predicting U.S. elections compared with polls at the time, but evidence suggests that they add little information when compared with modern polls (*40*).

To evaluate the contribution of polling data, we recreated the entire analysis process, starting from feature selection, without the polling variables (supplementary materials). The features selected in the models without the polling data were the same as those selected in models with the polling data. The performance on the training data was also acceptable, with about 75.5% accuracy (about 3.4% lower than the model that included polls) (table S4). The out-of-sample results, however, were not very strong, with about 64.8% accuracy, much worse than even the naïve model (table S5). This suggests that, without the polling data, the model is prone to overfitting. This is likely due to the large amount of heterogeneity in structural variables from a global sample over a long period of time and suggests a potential limiting factor in prediction from structural features alone.

Second, because of the long time frame covered, our polling variable was necessarily rough, often reflecting second-hand news reports about who was winning in polls versus the raw polling numbers. Could we improve predictions by better aggregating polling data?

To test this possibility, we crawled numerous media sources within Lexis-Nexis and Wikipedia to create a compendium of polling data for 146 rounds in 122 elections. In sum, we collected information from 4331 different polls that were released publicly. We found that even with spotty polling data, in which elections had fewer than five publicly available polls, polling data still had an outsized effect on predicting global presidential elections compared to other variables (figs. S6 and S7). As we demonstrate below, however, this polling data can be substantially improved by modeling potential biases.

We used these data to predict the margin of victory or defeat of the incumbent party's candidate—the number of percentage points won by the incumbent party candidate minus that of the highest-scoring opposition candidate. This measure was comparable across countries, including situations where there were more than two candidates and where runoff elections are required (supplementary materials). Here again, we excluded cases in which there is no incumbent party candidate.

As with the previous model, we incorporated several structural variables. These included the percentage change in the GDP in the current year from the 10-year average, the inflation rate, level of democracy, the prior bias of polling houses, and whether an incumbent officeholder was running (tables S8 and S9). In addition, our model incorporated the timeliness of polling

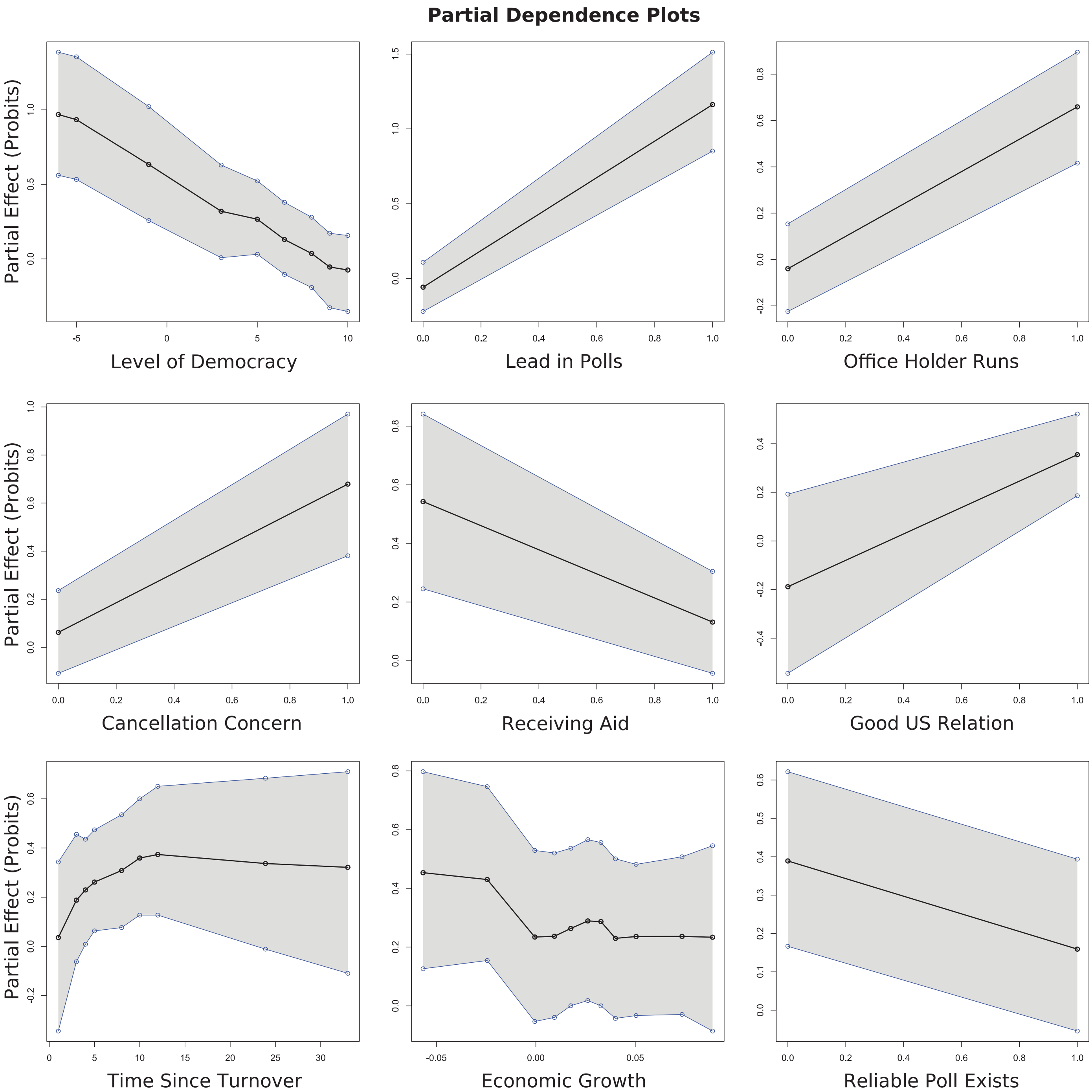

Fig. 1. Partial dependence plots from baseline model. Partial dependence, the effect of a variable with other variables held constant, was calculated as the marginal impact on the probit probability of the incumbent party winning at quantiles indicated by the dots on the lines. The 95% confidence intervals are calculated from the posterior distribution of Markov chain Monte Carlo analyses.

information to weight our predictions toward more timely polls.

We divided the data into an in-sample training set of 50 election rounds and an out-of-sample testing set of 96 rounds. In these elections, a naïve model that chose the incumbent party candidate to win in every case achieved 62.3% accuracy. Similarly, when we included only the slow-moving structural variables relating to institutions and the economy, we achieved a level of accuracy comparable to that of the naïve model, with an overall accuracy of 63.5% and an out-of-sample root mean squared error (RMSE) of 22.1 (fig. S8). Although this structural model did show a relationship between democracy and incumbent party vote share, economic growth showed no reliable relationship (supplementary materials).

Another baseline model used a median of the available polls to decide the winner (*41*), which yielded a correct prediction 86.3% across the entire sample—87.9% correct if the incumbent officeholder was running, 85.2% in the case of a chosen successor and a RMSE of 13.13. The polls-only model yielded an out-of-sample accuracy of 88.5% and a RMSE of 11.96. Although this was better than the accuracy of our earlier model that used a rougher measure of public opinion, there was still systematic measurement error due to elections with sparse data and polling house effects—some organizations had a pro- or anti-incumbent party bias. In addition to polling house effects, weaknesses in polling data were likely to be correlated with country, region, and election-specific factors, owing to diverse historical, institutional, and cultural circumstances.

To deal with these issues, we used model-based poll aggregation to create a "smoothed" polling estimate for public opinion rather than the polls themselves. This approach drew on country-, region-, and election-specific information to compensate for weaknesses in the polling data (supplementary materials). This model was fitted in three steps. Step 1 created a smoothed estimate of polls with a linear mixed-effects model. Step 2 created a prediction of the expected polling margins from step 1. Step 3 aggregated the data, taking the median prediction from the smoothed estimate. That smoothed polling estimate was then used alongside democracy scores and inflation rates in a partial least squares (PLS) model to forecast the eventual margin of the incumbent in each election.

This strategy was more accurate than the poll averages alone and far better than a purely structural model. Figure 2A displays the predicted margin of the polls for the incumbent party candidate on the x axis and the ultimate margin of the incumbent on the y axis. The points were colored to indicate different election rounds. The polls tracked tightly to the ultimate margin of the incumbent after smoothing. If one compares the raw polling data to the smoothed data, the effect is even more obvious (fig. S8). This demonstrates the advantage of global models—information from seemingly unrelated elections can help improve prediction across a variety of states.

After training the model with the first 50 election rounds, we then resampled and cross-validated using a rolling election window. We achieved over 90% accuracy in forecasting elections out of sample (86.0% in initial sample, 92.7% out of sample). The out-of-sample RMSE was estimated as 11.93 when the same push-forward resampling was used. The improvement in RMSE over the polls-only model was mild compared to the improvement in accuracy. This implies that the smoothing model makes the strongest contribution in cases where the raw polling data point to a tight race. The smoothing model added additional information to resolve measurement error in the polling data. The out-of-sample R^2—the proportion of margin variance captured by the model—was 0.74. In Fig. 2B we plotted the prediction from this model against the real margin of the incumbent party candidate, with smoother lines overlaid on the correct and incorrect predictions. In Fig. 2, C and D, we mapped and listed the correct and incorrect predictions from the model. As the sample grew from the initial 50 training election rounds, the accuracy of the model increased, suggesting that statistical regularities in elections across time and space are aiding the learning process (figs. S10 to S12). Predictions based on polls available weeks before the election were between 86 and 94% accurate (fig. S14).

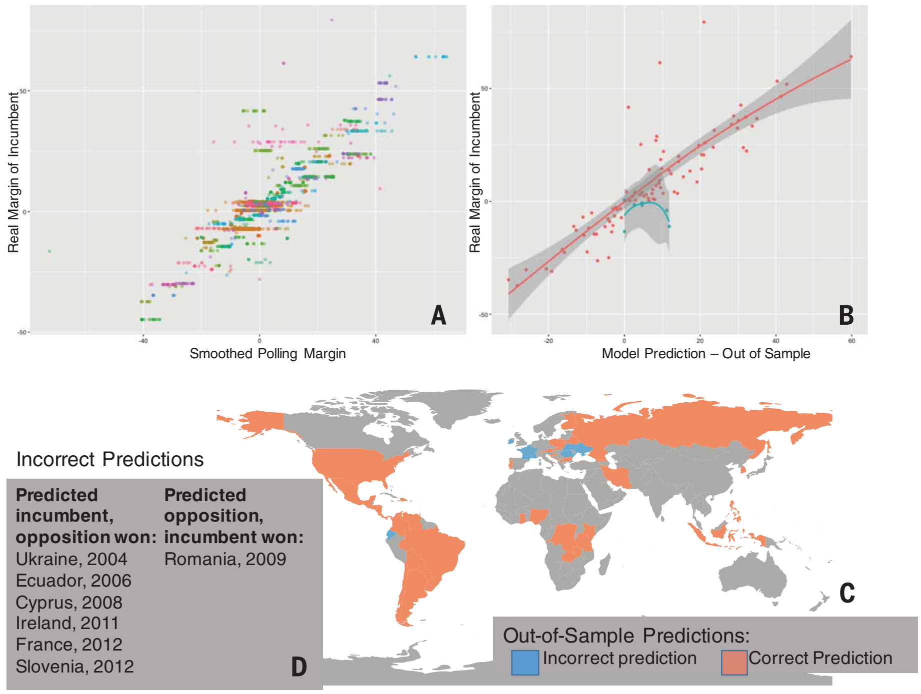

Fig. 2. Summary of results from the polling model. (**A**) Predictions of polling margins from the smoothing model are plotted against the ultimate margin of the incumbent party candidate, with different colors indicating different elections. (**B**) The final predictions from the structure + polling model are plotted against the ultimate margin of the incumbent. Smoother lines are overlaid, with red indicating correct predictions and blue, incorrect predictions. (**C**) Maps show countries in which a correct prediction was produced (in red), and the countries in which an incorrect prediction was made from the polling model (in blue). (**D**) List of the countries in which incorrect predictions were made from the polling model.

Much like other scholars (*25*, *42*), we uncovered substantial polling house effects, in which certain polling houses reliably supported the incumbent whereas others reliably supported the opposition. The most pro-opposition pollster has underestimated the margin for the incumbent by about 12 points, whereas the most pro-incumbent pollsters, including a few from the United States, overestimated by about 13 points (Fig. 3). Such consistent house effects allow for statistical adjustments to mitigate pro-opposition or pro-incumbent biases (tables S8 and S9).

Finally, from a structural perspective, although there is a long literature on the effect of economic growth in elections, we found little to suggest a global rule, with only minor impacts for inflation observed (table S10 and figs. S15 and S16). Less democratic institutions, unsurprisingly, tended to favor the incumbent party. We also looked at the effect of multiparty versus two-party elections, but found no characteristics that were consistently predictive of incumbent party margins (fig. S13).

In conclusion, we set out to test whether global models of election prediction could be successful. Overall, the results of our attempts at this were encouraging. Using relatively rough data over more than 70 years, we could predict the outcome of global elections with ~80% accuracy, including when we conducted live predictions in real time. This accuracy was increased to more than 90% when we were able to incorporate more detailed information about public opinion and the sources of polls. This increase in accuracy may be surprising to some, given that we were using data across countries with very different voting and party systems, and different levels of experience. Thus, we have shown that a global election prediction system is viable, and such efforts are likely to become more successful in the future, as the Internet continues to expand access to information on elections around the world and polling techniques become more widespread. Others may be surprised that we did not find a larger impact of economic variables, suggesting that the effects of economic factors are moderated by country-specific factors.

These results are given greater salience with the recent U.S. presidential election outcome. Few quantitative analysts predicted that Donald Trump would win. For our part, we presented a prediction of the election at the 2016 American Political Science Association (APSA) conference that gave Trump only a 16% chance of winning. Despite our prediction being technically based on popular vote totals, which Hillary Clinton did win (by 2%), the outcome was still much closer than we had predicted (about 7%) (fig. S20).

In response, some in the media have called into question the value of these quantitative forecasts. Reactions included: "Tonight data died" (*43*); "We should all feel bamboozled" (*44*); and "The pollsters…won't be believed on anything soon" (*45*). These reactions echo the reactions after the polling misses in the 1948 U.S. presidential election.

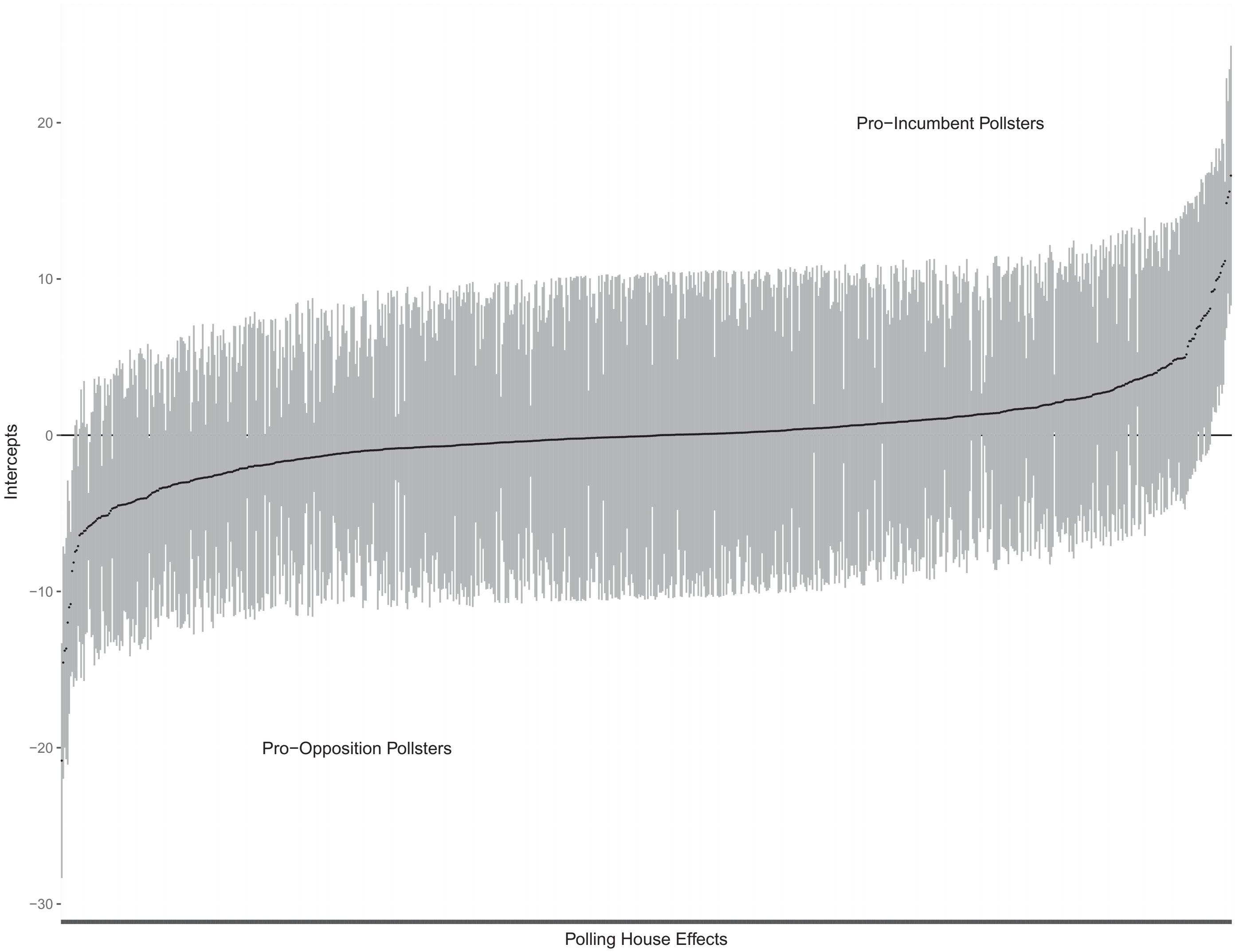

Fig. 3. Polling house bias in structure + polling model. The plot displays the estimated intercepts for the margin of the incumbent candidate for different polling houses. These polling house effects are ordered from the most pro-opposition to the most pro-incumbent.

As the *New York Times* reported the day after Truman won that election: "The polls were unable to compute statistically the unpredictable and unfathomable nuances of human character" (*46*). Our results highlight that neither public opinion nor structural factors, like economic growth, are ever going to be perfect predictors of election outcomes. They can, however, provide a generally accurate representation of likely election outcomes and help us overcome the many biases associated with human "gut feelings." We predict that reports of the death of quantitative electoral forecasts are greatly exaggerated.

REFERENCES AND NOTES

1. R. Fair, *Predicting Presidential Elections and Other Things* (Stanford Univ. Press, ed. 2, 2011).
2. J. E. Campbell, *PS Polit. Sci. Polit.* **20**, 37–42 (1987).
3. J. E. Campbell, *PS Polit. Sci. Polit.* **37**, 733–735 (2004).
4. J. E. Campbell, *PS Polit. Sci. Polit.* **47**, 769–771 (2014).
5. B. Lockerbie, *PS Polit. Sci. Polit.* **41**, 713–716 (2008).
6. P. Bellucci, *Int. J. Forecast.* **26**, 54–67 (2010).
7. A. Ceron, L. Curini, S. M. Iacus, G. Porro, *New Media Soc.* **16**, 340–358 (2014).
8. R. Nadeau, M. S. Lewis-Beck, É. Bélanger, *Int. J. Forecast.* **26**, 11–18 (2010).
9. M. S. Lewis-Beck, É. Bélanger, C. Fauvelle-Aymar, *Fr. Polit.* **6**, 106–115 (2008).
10. P. C. Magalhães, L. Aguiar-Conraria, M. S. Lewis-Beck, *Int. J. Forecast.* **28**, 769–776 (2012).
11. P. F. Whiteley, *Br. J. Polit. Int. Relat.* **7**, 165–173 (2005).
12. S. D. Fisher, R. Ford, W. Jennings, M. Pickup, C. Wlezien, *Elect. Stud.* **30**, 250–257 (2011).
13. E. T. K. Sang, J. Bos, in *Proceedings of the Workshop on Semantic Analysis in Social Media* (Association for Computational Linguistics, Stroudsburg, PA, 2012), pp. 53–60; http://dl.acm.org/citation.cfm?id=2389969.2389976.
14. B. Jérôme, V. Jérôme-Speziari, M. S. Lewis-Beck, *PS Polit. Sci. Polit.* **46**, 479–480 (2013).
15. A. Graefe, *Ger. Polit.* **24**, 195–204 (2015).
16. J. Ulfelder, Why the world can't have a Nate Silver, *Foreign Policy* (2012); https://foreignpolicy.com/2012/11/08/why-the-world-cant-have-a-nate-silver/.
17. M. Ward, N. Metternich, Predicting the future is easier than it looks, *Foreign Policy* (2012); https://foreignpolicy.com/2012/11/16/predicting-the-future-is-easier-than-it-looks/.
18. M. S. Lewis-Beck, M. Stegmaier, *PS Polit. Sci. Polit.* **47**, 284–288 (2014).
19. A. Abramowitz, *PS Polit. Sci. Polit.* **45**, 618–619 (2012).
20. J. E. Campbell, *PS Polit. Sci. Polit.* **45**, 610–613 (2012).
21. H. Norpoth, *PS Polit. Sci. Polit.* **47**, 332–335 (2014).
22. J. Berg, R. Forsythe, F. Nelson, T. Rietz, *Handb. Exp. Econ. Results* **1**, 742–751 (2008).
23. D. Rothschild, *Int. J. Forecast.* **31**, 952–964 (2015).
24. C. E. Cook, D. Wasserman, *PS Polit. Sci. Polit.* **47**, 304–308 (2014).
25. D. A. Linzer, *J. Am. Stat. Assoc.* **108**, 124–134 (2013).
26. S. Jackman, *PS Polit. Sci. Polit.* **47**, 317–321 (2014).
27. R. Nadeau, M. S. Lewis-Beck, É. Bélanger, *Comp. Polit. Stud.* **46**, 551–573 (2013).
28. R. M. Duch, R. T. Stevenson, *The Economic Vote: How Political and Economic Institutions Condition Election Results* (Cambridge Univ. Press, 2008).
29. J. Ufelder, *Foreign Policy* (8 November 2012).
30. S. D. Hyde, N. Marinov, *Polit. Anal.* **20**, 191–210 (2012).
31. H. A. Chipman, E. I. George, R. E. McCulloch, *Ann. Appl. Stat.* **4**, 266–298 (2010).
32. A. Kapelner, J. Bleich, http://arxiv.org/abs/1312.2171 (2013).
33. J. Bleich, A. Kapelner, E. I. George, S. T. Jensen, *Ann. Appl. Stat.* **8**, 1750–1781 (2014).
34. A. Gelman, G. King, *Am. J. Pol. Sci.* **34**, 1142–1164 (1990).
35. S. Ansolabehere, J. M. Snyder, *Election Law J. Rules Polit. Policy* **1**, 315–338 (2002).
36. S. Knack, *Int. Stud. Q.* **48**, 251–266 (2004).

37. C. C. Gibson, B. D. Hoffman, R. S. Jablonski, *World Dev.* **68**, 323–335 (2015).
38. A. A. Licht, *J. Conflict Resolut.* **54**, 58–87 (2010).
39. P. Norris, R. W. Frank, F. M. Coma, *J. Democracy* **24**, 124–135 (2013).
40. R. S. Erikson, C. Wlezien, *Elect. Stud.* **31**, 532–539 (2012).
41. J. Richard Gott III, W. N. Colley, *Math. Comput. Model.* **48**, 1396–1408 (2008).
42. F. Cantú, V. Hoyo, M. A. Morales, *Int. J. Public Opin. Res.* **28**, 96–116 (2015).
43. J. Rutenberg, "A 'Dewey Defeats Truman' Lesson for the Digital Age," *New York Times*, 9 November 2016; www.nytimes.com/2016/11/09/business/media/media-trump-clinton.html.
44. "Our fictional pundit predicted more correct primary results than Nate Silver did," *Washington Post*, 9 May 2016; www.washingtonpost.com/posteverything/wp/2016/05/09/our-fictional-pundit-predicted-more-correct-primary-results-than-nate-silver-did/.
45. K. P. Vogel, A. Isenstadt, "How did everyone get it so wrong?" *POLITICO*, 9 November 2016; http://politi.co/2feuVaZ.
46. C. Phillips, "And the polls: What went wrong? Sampling technique is apparently not so 'scientific,'" *New York Times*, 7 November 1948, p. E4.

ACKNOWLEDGMENTS

Supported in part by the Intelligence Advanced Research Projects Activity (IARPA) through DoI/NBC contract no. D12PC000337. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the U.S. government. We acknowledge the invaluable work of the research assistants who helped with data collection in all parts of this project.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/355/6324/515/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S19
Tables S1 to S10
References (*47*–*61*)

26 October 2016; accepted 24 December 2016
10.1126/science.aal2887

Improving election prediction internationally
Ryan Kennedy, Stefan Wojcik and David Lazer (February 2, 2017)
Science **355** (6324), 515-520. [doi: 10.1126/science.aal2887]

Editor's Summary

Using global data for election predictions

Assumptions underlying election result predictions have been questioned recently. Kennedy *et al.* assessed more than 650 executive office elections in over 85 countries and performed two live forecasting experiments. They analyzed a variety of potential predictors theorized to be of importance, ranging from economic performance to polling data. Elections were about 80 to 90% predictable, despite uncertainties with available data. Polling data were very important to successful prediction, although it was necessary to correct for systematic biases. Unexpectedly, economic indicators were only weakly predictive. As data sources improve and grow, predictive power is expected to increase.

Science, this issue p. 515