POLITICS

Special Issue on Forecasting the 2024 US Elections

Lessons Learned: Citizen Forecasting, Candidate Resignations, and the 2024 US Presidential Election

Brian Thompson-Collart, *Université Laval, Canada*
Hubert Cadieux, *Université Laval, Canada*
Catherine Ouellet, *Université de Montréal, Canada*
Yannick Dufresne, *Université Laval, Canada*

Abstract

Every four years, a slew of election forecasting models attempt to predict the results of the US presidential election. Regardless of the stability of any election system, such as the bipartisan system in the US, conditions can arise (e.g., candidate resignations) that negatively impact forecasters' ability to predict electoral outcomes. Citizen forecasting, or directly asking respondents who will win an election, enjoys a long track record in successfully predicting presidential elections. This article proposes adapting a citizen forecasting measure originally intended for use in multiparty systems to predict the US presidential election in 2024. Using this measure, we create a forecast of the national-level popular vote and vote share forecasts for seven swing states.

This is a "preproof" accepted article for PS: Political Science & Politics. This version may be subject to change during the production process.

DOI: 10.1017/S1049096524000969

Introduction

The results of future elections matter a great deal to different groups of stakeholders (e.g., voters, policymakers, and elected representatives). Researchers have developed several types of models to forecast elections. Currently, these researchers generally employ three types of models to forecast US presidential elections (Murr and Lewis-Beck 2020, 91). The first type, econometric models, employ aggregate-level data and regression techniques to estimate incumbents' vote or seat shares (Bélanger and Trotter 2017, 821). These models assume that the electorate rewards or punishes incumbents based on economic performance. Second, researchers employ prediction markets where traders can buy and sell contracts corresponding to real-life election outcomes. (Luckner 2012, 7). The third type of election forecasting model analyzes individual-level responses to vote intention and vote expectation items on survey instruments. Vote intention items ask respondents what candidate or party they intend to vote for in an upcoming election. Vote expectation items ask respondents what candidate or party they think will win an upcoming election (Rothschild and Wolfers 2012, 4). Although researchers have employed citizen forecasting extensively in US presidential elections, no study has employed a likelihood citizen forecasting measure to predict election outcomes within that context.

This paper makes three contributions to the literature. First, this paper represents the first use of a likelihood citizen forecasting measure in US presidential elections, a generally bipartisan system. Murr (2011) used this type of measure to aggregate citizen forecasts in the British multiparty system. Although most US presidential elections pit candidates from two major parties against each other, sometimes strong third-party candidates appear in bipartisan

systems. Before ending his presidential campaign on August 23, 2024, Robert F. Kennedy Jr. appeared on the ballot in at least 23 states and sought ballot access in 25 others (Slisco, 2024). Although Kennedy ended his campaign and endorsed Donald Trump, his campaign's rise shows the need to employ citizen forecasting measures that can capture levels of support for more than two candidates. Collecting citizen forecasts using a likelihood measure possesses several advantages over a categorical measure. For example, we can directly observe the level of (un)certainty that citizens have in their forecasts at the national and state level. In addition, we can estimate vote shares more easily. Previous research generally uses historical data to predict a party's vote share in US elections (Murr 2015, 922). Asking citizens to rate a candidate's chances using a likelihood measure represents an improvement over this process. It allows us to estimate a vote share forecast for each candidate in the election without resorting to historical data.

The paper's second contribution includes a forecast of the share of the national popular vote that citizens expect Joe Biden and Donald Trump to receive. We caveat this forecast and our swing state forecast by emphasizing that both forecasts only apply to the state of the race before July 2024. On July 21, 2024, Joe Biden announced that he would not stand for re-election and endorsed Vice-President Kamala Harris to replace him on the Democratic ticket (Baker 2024). Biden ending his re-election campaign, Harris entering the race, and RFK Jr ending his campaign a month later changed the race to the extent that forecasts asking about these candidates' chances of victory should not be assessed retrospectively. However, in this article we wish to show how a likelihood measure can successfully be applied at the national and state levels in future US presidential elections and what lessons future election forecasters can learn from this highly volatile election.

The paper's third contribution consists of a forecast for seven states we identify as swing states. We present vote share forecasts for the states of Florida, Georgia, Michigan,

North Carolina, Ohio, Pennsylvania, and Virginia. The researchers chose to focus on these states as swing states for three reasons. First, the Electoral College renders most states uncompetitive in US presidential elections. Previous citizen forecasting studies recommend focusing on competitive electoral districts within a country over uncompetitive ones as these represent a more stringent test of citizen forecasting (Thompson-Collart et al. 2024, 8). Second, the states in our sample demonstrate competitiveness in recent elections. The winning candidate carried five of seven states in our sample by five points or less in the 2016 and 2020 elections (Wolf 2024). Third, polling aggregators identify four of the states in our sample---Georgia, Michigan, North Carolina, and Pennsylvania, as particularly consequential states that could 'tip' the election, (i.e., push a candidate over the 270 electoral votes needed to win the presidency; FiveThirtyEight 2024). As of September, polling aggregator FiveThirtyEight gave Pennsylvania a 17.5% probability of tipping the election, North Carolina receives a 12% probability, Georgia 11.5%, and Michigan 11.4%. For these reasons, we choose to limit our sample to these states.

Citizen Forecasting Methodology

Citizen forecasting aggregates individual predictions to provide a forecast of what candidate or party will win an upcoming election. This technique relies on Condorcet's jury theorem (Murr 2011, 771; Murr 2015, 917; Temporão et al. 2019, 3). Under Condorcet's original formulation, citizens had to each possess a greater than fifty percent probability of making a correct prediction, their votes had to be uncorrelated, and the predicted outcome had to be binary (Murr 2011, 772). Subsequent research relaxes these assumptions. These relaxed assumptions allow both the competence levels of citizens and the correlation of votes to vary as well as allow for predictions with multiple outcomes (Murr 2015, 918). Citizens' competence levels represent the key to a successful citizen forecast. If a group of citizens has a greater than even chance of predicting the correct outcome, the probability of making a

correct election forecast approaches 100% as citizens are added to the group (Murr 2015, 917). Whether unrepresentative samples of citizens can predict election results remains an open research question. Previous studies show that unrepresentative, but highly competent, samples of citizens within US states can usually predict the presidential election in their state (Murr 2015, 919). However, other studies find that unrepresentative samples do not outperform a representative sample (Ganser and Riordan 2015, 124).

Citizen forecasting studies use two methods for aggregating citizens' predictions about upcoming elections. The first of these methods, plurality voting, tallies the percentage of respondents that believe a specific party will win an election (Murr 2011, 774). The party with the highest proportion of individuals expecting that party to win is forecast as the election winner. Although plurality voting is relatively straightforward, this method discards a considerable amount of information, such as what party came in second or third as well as the level of certainty each respondent has in their forecast. Murr (2011) proposes range voting as an alternative method to plurality voting in multiparty elections. Range voting sums and normalises expectation scores from a likelihood measure (Murr 2011, 774). Range voting provides two advantages over plurality voting. First, this method provides information on what parties will come in second and third places. Second, this method allows analysts to observe individual respondent's level of certainty in their election prediction.

This article aggregates citizens' forecasts about the 2024 US presidential election using a range voting procedure. We asked citizens what candidate they think will win the election at the national level and in their state. The national-level question asks respondents, "How likely do you think it is that Donald Trump, Joe Biden, or RFK Jr. will be elected president in November? Please assign a probability to each candidate." The response options included Donald Trump, Joe Biden, and Robert F. Kennedy Jr. The state-level question asked respondents, "How likely do you think it is that Donald Trump, Joe Biden, or RFK Jr. will

win your state in the presidential election? Please assign a probability to each candidate." The response options again included Donald Trump, Joe Biden, and Robert F. Kennedy Jr. Respondents could assign a probability of victory to each candidate ranging from 0% to 100%. These measures allow analysts to not only predict the election winner but also the closeness of an election (Temporão et al. 2019, 4). These items produced two pieces of information needed to predict the outcome of a presidential election. First, this measure provides an average likelihood figure for each candidate. We predict that the candidate with the highest average likelihood will win the election. Second, we can use data collected using this measure to estimate vote shares for each of the candidates in the election. We obtain vote shares for all candidates in two manners: First, we divide the average likelihood for each candidate by the sum of all likelihoods for all the candidates. Second, we repeat this process using the median likelihood. We estimate vote shares using both mean and median likelihoods because the likelihood distributions for all three candidates are not symmetrical. As a result, taking the mean only might lead us to overestimate support for a minor candidate while underestimating support for the major candidates (Penn State 2024). The use of median likelihoods addresses this concern.

A practical example will serve to illustrate the vote share estimation process. To normalise the likelihood scores for each candidate, we first take the average likelihood score for each. For example, in the state of Michigan, the average likelihood score for Joe Biden was 3.6, the average likelihood score for Donald Trump was 5.5 and the average likelihood score for Kennedy was 1.8. To estimate the vote share for Joe Biden we divide his average likelihood of 3.6 by 10.9, or the sum of the likelihood scores for all the candidates (3.6 + 5.5 + 1.8 = 10.9). This procedure results in a vote share forecast of 33% for Joe Biden in Michigan. We estimate vote shares in this manner at the national level as well as within each state. We repeat the process using the median likelihood as well.

This technique for estimating vote shares from citizen forecasts represents an improvement compared to previous methods. To obtain vote shares from citizen forecasts, researchers usually regress the percentage of citizens that believe a party will win on the vote share obtained by that party in an election (Lewis-Beck and Tien 1999, 181; Murr 2011, 777). Although this method provides accurate results, it also suffers from pragmatic limitations. First, we require historical data to estimate a regression equation. Although historical data exists for the two major parties, we do not have historical data for new, third-party candidates that may crop up. Second, estimating vote shares in this way is less time-consuming than collecting historical data and estimating a regression. Finally, this method for estimating vote shares appears to provide accurate predictions. The technique demonstrated a mean absolute error of 2 percentage points for the first round of the 2017 French presidential election and 1.5 points for the second round (Dufresne et al. 2022, 732). The method provides an error rate similar to using the final Gallup poll in an election campaign (Lewis-Beck and Tien 1999, 183).

Citizen Forecasting in United States Presidential Elections

Both vote intention polling and citizen forecasting vary in their accuracy for predicting election results. Lewis-Beck and Tien (1999) first compared the accuracy of citizen forecasting to vote intention polling. They found that both vote intention polling and citizen forecasting correctly predicted the winner in nine of eleven elections between 1956 and 1996. Moreover, when comparing vote share estimates obtained from citizen forecasts to those from vote intention polls, they found a similar mean average prediction error across the two forecasting methods. Graefe (2014) extended this line of research to other election forecasting methods. He compared the accuracy of citizen forecasting to vote intention polling, prediction markets, and quantitative models for US presidential elections. He found that citizen forecasting predicted vote shares similarly to quantitative models and better than

vote intention polling and prediction markets. From these results, we can conclude that citizen forecasting represents an accurate election forecasting method that can complement vote intention polling.

Lewis-Beck and Skalaban (1989) first demonstrated that US citizens could accurately predict presidential elections. Across eight presidential election years between 1956 and 1988, 69% of citizens correctly predicted the winning party. Lewis-Beck and Tien (1999) subsequently found that contextual factors explained why some citizens could better predict elections than others. These contextual factors include respondent's level of education, date of the interview, and whether the respondent expects a close election. However, partisan affiliation exerts a considerable effect on citizen forecasts as well (Lewis-Beck and Skalaban 1989, 149; Lewis-Beck and Tien 1999, 179; Mongrain 2021a, 11). Further, Dolan and Holbrook (2001) found that political knowledge improves citizen forecasts and attenuates the effects of wishful thinking at the state level but uncover no such attenuation at the national level. Taken together, citizen forecasting provides researchers with a valuable method for accurately predicting US presidential elections.

The 2024 US Presidential Election: Data and Methodology

The researchers elicited citizen forecasts from a probabilistically selected sample of adults residing in the United States. Léger polling, a private polling firm, collected the survey data between May 13 and July 2, 2024. The sample contained 1,607 respondents at the national level. Figure A1 shows the number of respondents answering the survey per day during the data collection period. The largest number of respondents in a single state, 164, came from California. By contrast, the state with the smallest number of respondents was Vermont with only three. At the national level, the sample mostly reflected the U.S. population on key demographic characteristics. For example, 38% of the sample possessed a college degree and 44% of the sample was female. Moving to the state level, we see that a

total of only 16 states possessed at least 30 respondents. Of these 16 states, we identified seven as swing states (i.e., Florida, Georgia, Michigan, North Carolina, Ohio, Pennsylvania, and Virginia) with at least 30 respondents. Florida contained the most respondents, 105. Virginia contained the least respondents, 41. Therefore, we expect our swing state forecasts to reflect an acceptable level of accuracy.

Who Will Win the 2024 United States Presidential Election?

This paper examines whether citizen forecasts can predict the winner of the national popular vote and the winner in seven key swing states. In this section, we present the results from our citizen forecasting model of the 2024 United States presidential election. We also compare those results to a publicly available citizen forecast conducted two months prior. Our citizen forecasting model produces two forecasts. First, a national-level forecast describes the percentage of the national popular vote that citizens expected Donald Trump and Joe Biden to win at the time of data collection. Second, we present a vote share forecast that identifies the party expected to win in seven competitive swing states. We first examine the share of the national popular vote share that we expect each candidate to receive. Figure 1 shows that citizens expected Donald Trump would receive 55% of the national popular vote and Biden would receive 45% of the national popular vote during the data collection period. If we employ the median, Trump receives 53% and Biden 47%. Our citizen forecast predicts similar results to the April 2024 Verasight MPSA Omnibus Study that forecasted Trump to win 50% of the national popular vote while Biden received 38% (Leiter and Lewis-Beck 2024).

Moving to the state-level forecast in Figure 2, we see that our citizen forecasters clearly expected a Trump victory in all seven swing states under study. Regardless of whether we calculate state-level vote shares using the average or the median likelihood, we can see that respondents across all seven swing states also expected Trump to garner a sizeable share

of the vote in their state. If we calculate vote shares employing the average likelihood, we see that citizens' forecasts of Trump's vote shares ranged from a minimum of 53% in Virginia to a maximum of 67% in Florida. If we calculate vote shares employing the median likelihood, we see that citizens' forecasts of Trump's vote shares ranged from a minimum of 50% in Pennsylvania---a tie with Biden---to a maximum of 80% in North Carolina.

Figure 1

Vote Share Expectations at the National Level

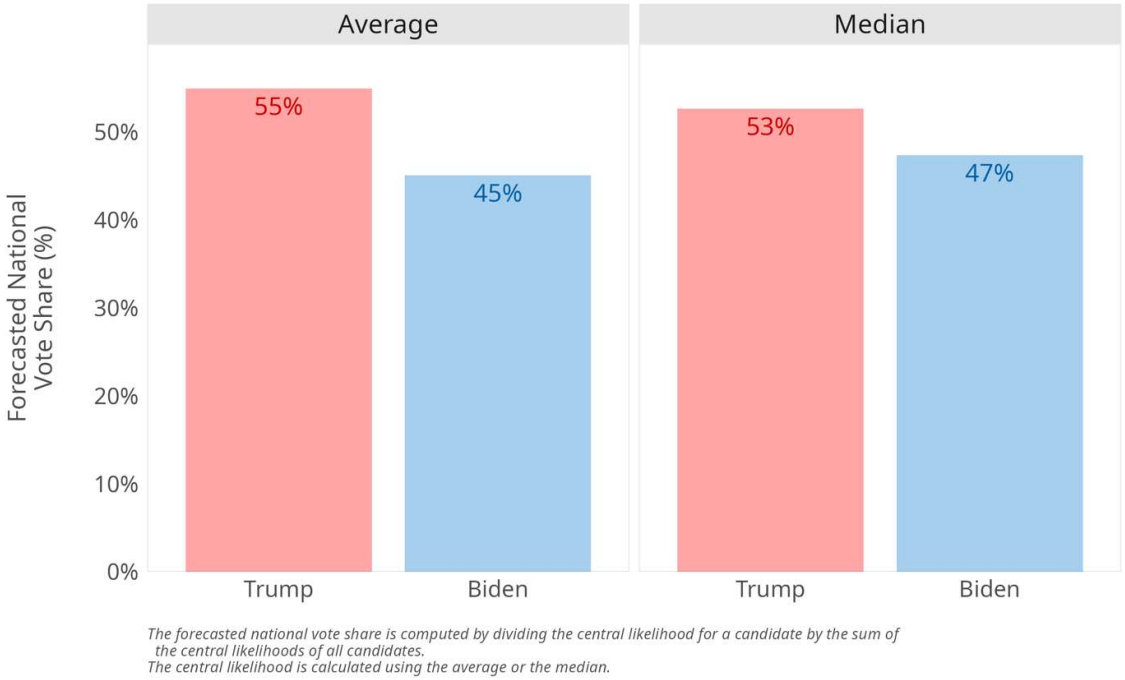
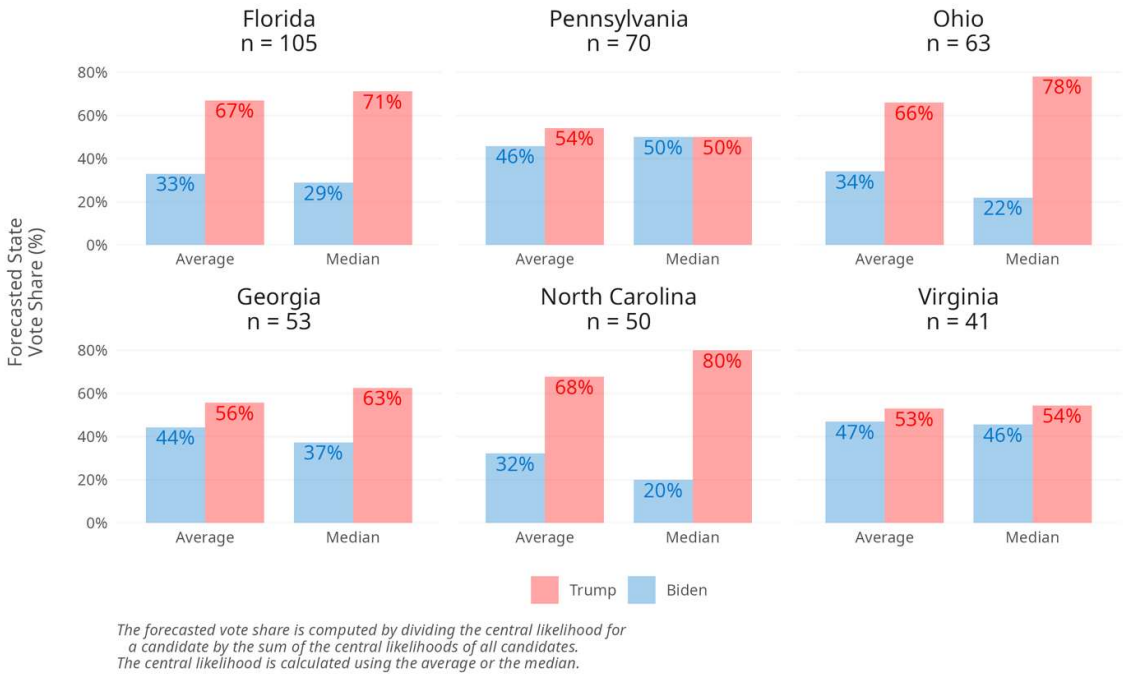


Figure 2

Vote Share Expectations in Selected Swing States



This project provided us with several lessons for conducting future citizen forecasts. First, future survey questions should ask citizens about *parties'* chances of winning an election, not *candidate's* chances. Although replacing candidates in the middle of a presidential campaign remains rare in United States presidential elections, the 2024 election showed that this can occur. Therefore, future researchers should ensure that their expectations measures ask about the likelihood that a party will win the presidency rather than a candidate. Second, future researchers should ensure that every state contains at least 30 respondents. Previous citizen forecasting studies have aimed to include this number of respondents because this reduces the uncertainty in state-level forecasts (Murr and Lewis-Beck 2020, 92). Our study contained at least 30 respondents in swing states but did not achieve this sample size in every state. Third, future researchers should seek to delegate national and state-level forecasts to the most competent citizens in the sample. To accomplish this, researchers should include items on future surveys measuring respondents' levels of political knowledge.

Previous research finds that delegating the citizen forecasting task to individuals possessing higher levels of political knowledge increases forecasting competence (Mongrain 2021b, 721).

Data Availability Statement

Research documentation and data that support the findings of this study have not yet been verified by PS's replication team. Data will be openly available at the Harvard Dataverse upon publication of the final article.

References

Baker, Graeme, and Michael S. McNamee. 2024. "Biden ends re-election bid, upending White House race." *BBC News.* July 21. https://www.bbc.com/news/articles/c1e5xpdzkd8o.

Bélanger, Érik. and Daniel Trotter. 2017. "Econometric approaches to forecasting." In *The SAGE Handbook of Electoral Behaviour*, edited by Kai Arzheimer, Jocelyn Evans, & Michael Lewis-Beck SAGE Publications Ltd.

Dolan, Kathleen, and Thomas Holbrook. 2001. "Knowing versus caring: The role of affect and cognition in political perceptions." *Political Psychology* 22(1): 27-44. https://doi.org/10.1111/0162-895X.00224.

Dufresne, Yannick, Bruno Jerome, Michael Lewis-Beck, Andreas Murr, and Justin Savoie. 2022. "Citizen forecasting: The 2022 French presidential election." *PS: Political Science and Politics* 55(4): 730-734. https://doi.org/10.1017/S1049096522000567.

Graefe, Andreas. 2014. "Accuracy of vote expectation surveys in forecasting elections." *Public Opinion Quarterly* 78(S1): 204-232. https://doi.org/10.1093/poq/nfu008.

FiveThirtyEight. September 20, 2024. "Who is favored to win the 2024 presidential election?" Accessed September 20, 2024. https://projects.fivethirtyeight.com/2024-election-forecast/.

Leiter, Debra, and Michael Lewis-Beck. 2024. "Citizen forecasting for the 2024 presidential race: A second sounding." The Center for Politics. June 5. https://centerforpolitics.org/crystalball/citizen-forecasting-for-the-2024-presidential-race-a-second-sounding/.

Lewis-Beck, Michael, and Andrew Skalaban. 1989. "Citizen forecasting: Can voters see into the future?" *British Journal of Political Science* 19(1): 419-427. https://doi.org/10.1017/S000712340000538X.

Lewis-Beck, Michael and Charles Tien. 1999. "Voters as forecasters: A micromodel of election prediction." *International Journal of Forecasting* 15(2): 175-184.

Luckner, Stefan. 2012. *Prediction markets: Fundamentals, designs, and applications.* Weisbaden: Gabler Verlag.

Mongrain, Philippe. 2021a. "Did you see it coming? Explaining the accuracy of voter expectations for district and (sub)national election outcomes in multi-party systems." *Electoral Studies* 71(e102317). https://doi.org/10.1016/j.electstud.2021.102317.

Mongrain, Philippe. 2021b. "A technocratic view of election forecasting: Weighting citizens' forecasts according to competence." *International Journal of Public Opinion Research* 33(3): 713-723. https://doi.org/10.1093/ijpor/edab010.

Murr, Andreas. 2011. "Wisdom of the crowds? A decentralised election forecasting model that uses citizens' local expectations." *Electoral Studies* 30(4): 771-783. https://doi.org/10.1016/j.electstud.2011.07.005.

Murr, Andreas. 2015. "The wisdom of crowds: Applying Condorcet's jury theorem to forecasting US presidential elections." *International Journal of Forecasting* 31(3): 916-929. https://doi.org/10.1016/j.ijforecast.2014.12.002.

Murr, Andreas, and Michael Lewis-Beck. 2020. "Citizen forecasting: A state-by-state experiment." *PS: Political Science and Politics* 54(1): 91-95. https://doi:10.1017/S1049096520001456.

Penn State Eberley College of Science. n.d. "2.2.4.1 – Skewness and central tendency." Accessed July 13, 2024. https://online.stat.psu.edu/stat200/lesson/2/2.2/2.2.4/2.2.4.1

Slisco, Aila. 2024. "Robert F. Kennedy may still be on the ballot in 23 states." *Newsweek.* August 24. https://www.newsweek.com/robert-f-kennedy-jr-may-still-ballot-23-states-1943742.

Temporão, Mickael, Yannick Dufresne, Justin Savoie, and Clifton van der Linden. 2019. "Crowdsourcing the vote: New horizons in citizen forecasting." *International Journal of Forecasting* 35(1): 1-10. https://doi.org/10.1016/j.ijforecast.2018.07.011.

Thompson-Collart, Brian, Evelyne Brie, and Yannick Dufresne. 2024. "Can Latin American voters see the future? Citizen forecasting in Argentina." *Revista Latinoamericana de Opinión Pública* 13(e31348). https://doi.org/10.14201/rlop.31348.

Wolf, Zachary B. and Renée Rigdon. 2024. "Why an election for 330 million-plus people may still come down to so few votes." *CNN.* August 31. https://www.cnn.com/2024/07/20/politics/electoral-votes-swing-state-margins-dg/index.html.