Post-election Analysis of Presidential Election/Poll Data: Liars Continued to Lie But What Have Changed?

Jiming Jiang

University of California, Davis, USA

April 2021

This work is joint with Yuanyuan Li of University of California, Davis and Peter X.-K. Song of the University of Michigan

Outline

1. Introduction

- In many ways, the past two presidential elections (2016, 2020) of the United States have been like no others before.

1. Introduction

- In many ways, the past two presidential elections (2016, 2020) of the United States have been like no others before.
- Predicting outcomes of the U. S. election has attracted interests of many political pollsters over the years, including statisticians.

1. Introduction

- In many ways, the past two presidential elections (2016, 2020) of the United States have been like no others before.
- Predicting outcomes of the U. S. election has attracted interests of many political pollsters over the years, including statisticians.
- Most notably in recent years is FiveThirtyEight (538), whose website (https://fivethirtyeight.com/) publishes an extensive range of results on election models and projections, including presidential, senate, and house of representative elections.

1. Introduction

- In many ways, the past two presidential elections (2016, 2020) of the United States have been like no others before.
- Predicting outcomes of the U. S. election has attracted interests of many political pollsters over the years, including statisticians.
- Most notably in recent years is FiveThirtyEight (538), whose website (https://fivethirtyeight.com/) publishes an extensive range of results on election models and projections, including presidential, senate, and house of representative elections.
- However, like many other prediction models, the 538 model also incorrectly predicted the 2016 USPE outcome.

1. Introduction

- In many ways, the past two presidential elections (2016, 2020) of the United States have been like no others before.
- Predicting outcomes of the U. S. election has attracted interests of many political pollsters over the years, including statisticians.
- Most notably in recent years is FiveThirtyEight (538), whose website (https://fivethirtyeight.com/) publishes an extensive range of results on election models and projections, including presidential, senate, and house of representative elections.
- However, like many other prediction models, the 538 model also incorrectly predicted the 2016 USPE outcome.
- To the best of our knowledge, there have been no academic studies using data science tools to jointly analyze the 2016 and 2020 elections.

- Yet, the joint analysis seems reasonable because both elections involved a (very) special candidate, Donald Trump.

- Yet, the joint analysis seems reasonable because both elections involved a (very) special candidate, Donald Trump.
- On the other hand, there have been informally published academic studies on the 2016 election (e.g., https://science.nu/community/the-great-society-reagans-revolution-and-generations-of-presidential-voting/)

- Yet, the joint analysis seems reasonable because both elections involved a (very) special candidate, Donald Trump.
- On the other hand, there have been informally published academic studies on the 2016 election (e.g., https://science.nu/community/the-great-society-reagans-revolution-and-generations-of-presidential-voting/)
- There have also been statistical analysis of election data concerning political elections elsewhere in the world (e.g., Brisco and Migliorati 2020).

- Yet, the joint analysis seems reasonable because both elections involved a (very) special candidate, Donald Trump.
- On the other hand, there have been informally published academic studies on the 2016 election (e.g., https://science.nu/community/the-great-society-reagans-revolution-and-generations-of-presidential-voting/)
- There have also been statistical analysis of election data concerning political elections elsewhere in the world (e.g., Brisco and Migliorati 2020).
- However, due to the unique natures of the USPE, namely, the EC system, and the special candidate, such results and models from other studies are less likely to be applicable to our problems of interest.

- For the most part, the EC assigns each state a given number of EC votes, with larger states (in terms of population) receiving more EC votes.

- For the most part, the EC assigns each state a given number of EC votes, with larger states (in terms of population) receiving more EC votes.
- If a candidate wins a state by simple majority, the candidate collects all the EC votes attached to the state.

- For the most part, the EC assigns each state a given number of EC votes, with larger states (in terms of population) receiving more EC votes.
- If a candidate wins a state by simple majority, the candidate collects all the EC votes attached to the state.
- Number of EC votes then add up over the 50 states to yield the total of EC votes, for each candidate, and whoever receives 270 or more of the EC votes wins the presidency.

- For the most part, the EC assigns each state a given number of EC votes, with larger states (in terms of population) receiving more EC votes.
- If a candidate wins a state by simple majority, the candidate collects all the EC votes attached to the state.
- Number of EC votes then add up over the 50 states to yield the total of EC votes, for each candidate, and whoever receives 270 or more of the EC votes wins the presidency.
- Due to this special feature of the EC, it makes sense to think of a state-based prediction problem.

2. SAE Models for Election Data

- To address such a problem, small area estimation (SAE; e.g., Rao and Molina 2015) offers a nearly ideal platform, as we shall see.

2. SAE Models for Election Data

- To address such a problem, small area estimation (SAE; e.g., Rao and Molina 2015) offers a nearly ideal platform, as we shall see.
- The term “small area” typically refers to a population for which reliable statistics of interest cannot be produced due to certain limitations of the available data

2. SAE Models for Election Data

- To address such a problem, small area estimation (SAE; e.g., Rao and Molina 2015) offers a nearly ideal platform, as we shall see.
- The term "small area" typically refers to a population for which reliable statistics of interest cannot be produced due to certain limitations of the available data
- Examples of domains include a geographical region (e.g., a state, county or municipality), a demographic group (e.g., a specific age×sex×race group), a demographic group within a geographic region, and so on.

2. SAE Models for Election Data

- To address such a problem, small area estimation (SAE; e.g., Rao and Molina 2015) offers a nearly ideal platform, as we shall see.
- The term "small area" typically refers to a population for which reliable statistics of interest cannot be produced due to certain limitations of the available data
- Examples of domains include a geographical region (e.g., a state, county or municipality), a demographic group (e.g., a specific age×sex×race group), a demographic group within a geographic region, and so on.
- A basic strategy in SAE is called "borrowing strength". Basically, it says that one can do better than the small number of domain-wide observation by borrowing strength, or information, from other domains and/or sources.

- The borrowing of strength is carried out via utilizing a statistical model, typically a mixed-effects model. See, e.g., Jiang & Lahiri (2006), Rao & Molina (2015).

- The borrowing of strength is carried out via utilizing a statistical model, typically a mixed-effects model. See, e.g., Jiang & Lahiri (2006), Rao & Molina (2015).
- Here, a small area is formed by considering different factors of interest; for example, it may correspond to a state (including D.C.), state by party (Democrat/Republican), or state by party by pollster (e.g., SurveyMonkey), depending on the model involved in the analysis.

- The borrowing of strength is carried out via utilizing a statistical model, typically a mixed-effects model. See, e.g., Jiang & Lahiri (2006), Rao & Molina (2015).
- Here, a small area is formed by considering different factors of interest; for example, it may correspond to a state (including D.C.), state by party (Democrat/Republican), or state by party by pollster (e.g., SurveyMonkey), depending on the model involved in the analysis.
- Several SAE models for the election data are proposed and discussed in the sequel.

- The borrowing of strength is carried out via utilizing a statistical model, typically a mixed-effects model. See, e.g., Jiang & Lahiri (2006), Rao & Molina (2015).
- Here, a small area is formed by considering different factors of interest; for example, it may correspond to a state (including D.C.), state by party (Democrat/Republican), or state by party by pollster (e.g., SurveyMonkey), depending on the model involved in the analysis.
- Several SAE models for the election data are proposed and discussed in the sequel.
- It turns out that, regardless of a specific model being used, the overall conclusion is consistent.

- The borrowing of strength is carried out via utilizing a statistical model, typically a mixed-effects model. See, e.g., Jiang & Lahiri (2006), Rao & Molina (2015).
- Here, a small area is formed by considering different factors of interest; for example, it may correspond to a state (including D.C.), state by party (Democrat/Republican), or state by party by pollster (e.g., SurveyMonkey), depending on the model involved in the analysis.
- Several SAE models for the election data are proposed and discussed in the sequel.
- It turns out that, regardless of a specific model being used, the overall conclusion is consistent.
- The conclusion is ...

- A potential reason for the nearly total failure of the 2016 polls is that, as many have suspected, a significant proportion of voters for the Republican candidate simply did not tell the truth to the poll surveys, the so-called dishonest respondents, or "liars".

- A potential reason for the nearly total failure of the 2016 polls is that, as many have suspected, a significant proportion of voters for the Republican candidate simply did not tell the truth to the poll surveys, the so-called dishonest respondents, or "liars".
- There are possible reasons for such a behavior but this is beyond our focus of this paper for investigating.

- A potential reason for the nearly total failure of the 2016 polls is that, as many have suspected, a significant proportion of voters for the Republican candidate simply did not tell the truth to the poll surveys, the so-called dishonest respondents, or “liars”.
- There are possible reasons for such a behavior but this is beyond our focus of this paper for investigating.
- Have these behaviors been changed since 2016, given that the Republican candidate was the sitting President during the 2020 election?

- A potential reason for the nearly total failure of the 2016 polls is that, as many have suspected, a significant proportion of voters for the Republican candidate simply did not tell the truth to the poll surveys, the so-called dishonest respondents, or "liars".
- There are possible reasons for such a behavior but this is beyond our focus of this paper for investigating.
- Have these behaviors been changed since 2016, given that the Republican candidate was the sitting President during the 2020 election?
- If not, have the 2020 pollsters figured out some way to correct the infamous "lies" to the surveys?

SAE models

- We begin with a simple model with random intercepts at the state level.

SAE models

- We begin with a simple model with random intercepts at the state level.
- Some more complex models will be considered subsequently.

SAE models

- We begin with a simple model with random intercepts at the state level.
- Some more complex models will be considered subsequently.
- The outcome variable is the difference in (natural) logarithms of the final-week poll result, in terms of proportion of support for a given candidate, and that of the real election result.

SAE models

- We begin with a simple model with random intercepts at the state level.
- Some more complex models will be considered subsequently.
- The outcome variable is the difference in (natural) logarithms of the final-week poll result, in terms of proportion of support for a given candidate, and that of the real election result.
- That is, $y_{ijkt} = \log(p_{ijkt}/\pi_{ikt}) = \log(p_{ijkt}) - \log(\pi_{ikt})$, where i for state, j for pollster, k for party, t for year.

SAE models

- We begin with a simple model with random intercepts at the state level.
- Some more complex models will be considered subsequently.
- The outcome variable is the difference in (natural) logarithms of the final-week poll result, in terms of proportion of support for a given candidate, and that of the real election result.
- That is, $y_{ijkt} = \log(p_{ijkt}/\pi_{ikt}) = \log(p_{ijkt}) - \log(\pi_{ikt})$, where i for state, j for pollster, k for party, t for year.
- Thus, $y_{ijkt} < 0$ (> 0) means that the poll under-projects (over-projects) the election result.

- Some initial, simpler ideas about the small-area specific random effects have been entertained.

- Some initial, simpler ideas about the small-area specific random effects have been entertained.
- For example, it might be though that a simplest model would just have a univariate random intercept at the state level.

- Some initial, simpler ideas about the small-area specific random effects have been entertained.
- For example, it might be though that a simplest model would just have a univariate random intercept at the state level.
- However, when such a model is fitted, the restricted maximum likelihood (REML; e.g., Jiang 2007, sec. 1.3.2) estimate of the variance of the random effect is found to be zero.

- Some initial, simpler ideas about the small-area specific random effects have been entertained.
- For example, it might be though that a simplest model would just have a univariate random intercept at the state level.
- However, when such a model is fitted, the restricted maximum likelihood (REML; e.g., Jiang 2007, sec. 1.3.2) estimate of the variance of the random effect is found to be zero.
- An alternative model is fitted, which assumes univariate and independent random effects at the state-candidate level.

- Some initial, simpler ideas about the small-area specific random effects have been entertained.
- For example, it might be though that a simplest model would just have a univariate random intercept at the state level.
- However, when such a model is fitted, the restricted maximum likelihood (REML; e.g., Jiang 2007, sec. 1.3.2) estimate of the variance of the random effect is found to be zero.
- An alternative model is fitted, which assumes univariate and independent random effects at the state-candidate level.
- However, the independent assumption seems unreasonable.

- Four SAE models for election data:

- Four SAE models for election data:
- 1. A simple model

$$y_{ia} = \beta_0 + \beta_1 I_{\text{can},ia} + \beta_2 I_{\text{year},ia} + \beta_3 I_{\text{can},ia} I_{\text{year},ia} + z_{ia} v_i + e_{ia}, \qquad (1)$$

$i = 1, \ldots, 51,\ a = 1, \ldots, n_i.$

- Four SAE models for election data:
- 1. A simple model

$$y_{ia} = \beta_0 + \beta_1 I_{\text{can},ia} + \beta_2 I_{\text{year},ia} + \beta_3 I_{\text{can},ia} I_{\text{year},ia} + z_{ia} v_i + e_{ia}, \tag{1}$$

$i = 1, \ldots, 51$, $a = 1, \ldots, n_i$.

- 2. State-time level random effect

$$y_{itb} = \beta_0 + \beta_1 I_{\text{can},itb} + \beta_2 I_{\text{year},itb} + \beta_3 I_{\text{can},itb} I_{\text{year},itb} + z_{itb} v_{it} + e_{itb}, \tag{2}$$

$i = 1, \ldots, 51$, $b = 1, \ldots, n_{it}$.

- 3. Random effects at state and pollster levels

$$\begin{aligned} y_{ipc} &= \beta_0 + \beta_1 I_{\mathrm{can},ipc} + \beta_2 I_{\mathrm{year},ipc} + \beta_3 I_{\mathrm{can},ipc} I_{\mathrm{year},ipc} + z_{ipc} v_i \\ &\quad + u_p + e_{ipc}, \end{aligned} \tag{3}$$

$i = 1, \dots, 51$, $p \in S_i$, where S_i is a subset of $\{1, \dots, 175\}$ and 175 is the total number of pollsters, and $c = 1, 2, 3, 4$ denotes the combination of indexes for the two parties and two years.

- 3. Random effects at state and pollster levels

$$
\begin{aligned}
y_{ipc} \;=\;& \beta_0 + \beta_1 I_{\mathrm{can},ipc} + \beta_2 I_{\mathrm{year},ipc} + \beta_3 I_{\mathrm{can},ipc} I_{\mathrm{year},ipc} + z_{ipc} v_i \\
& + u_p + e_{ipc},
\end{aligned} \tag{3}
$$

$i = 1, \ldots, 51$, $p \in S_i$, where S_i is a subset of $\{1, \ldots, 175\}$ and 175 is the total number of pollsters, and $c = 1, 2, 3, 4$ denotes the combination of indexes for the two parties and two years.

- 4. Random effects at state-time and pollster levels

$$
\begin{aligned}
y_{itpd} \;=\;& \beta_0 + \beta_1 I_{\mathrm{can},itpd} + \beta_2 I_{\mathrm{year},itpd} + \beta_3 I_{\mathrm{can},itpd} I_{\mathrm{year},itpd} + z_{itpd} v_{it} \\
& + u_p + e_{itpd},
\end{aligned} \tag{4}
$$

$i = 1, \ldots, 51$, $p \in S_i$.

- Estimation results: Model (1)

Parameter	Estimate	Standard Error	t-statistic	Group
β_0	-0.055	0.009	-5.903	Fixed
β_1	-0.081	0.020	-4.030	Fixed
β_2	0.072	0.006	12.057	Fixed
β_3	-0.018	0.009	-2.151	Fixed
σ_{d}	0.059	0.007	8.997	State
σ_{r}	0.082	0.009	9.427	State
ρ	-0.891	0.050	-17.704	State
τ	0.082	0.001	65.628	Residual

- Estimation results: Model (2)

Parameter	Estimate	Standard Error	t-statistic	Group
β_0	-0.060	0.009	-6.550	Fixed
β_1	-0.076	0.020	-3.865	Fixed
β_2	0.084	0.013	6.404	Fixed
β_3	-0.034	0.028	-1.186	Fixed
σ_{d}	0.056	0.006	10.036	Year:State
σ_{r}	0.079	0.007	11.731	Year:State
ρ	-0.890	0.043	-20.534	Year:State
τ	0.080	0.001	54.298	Residual

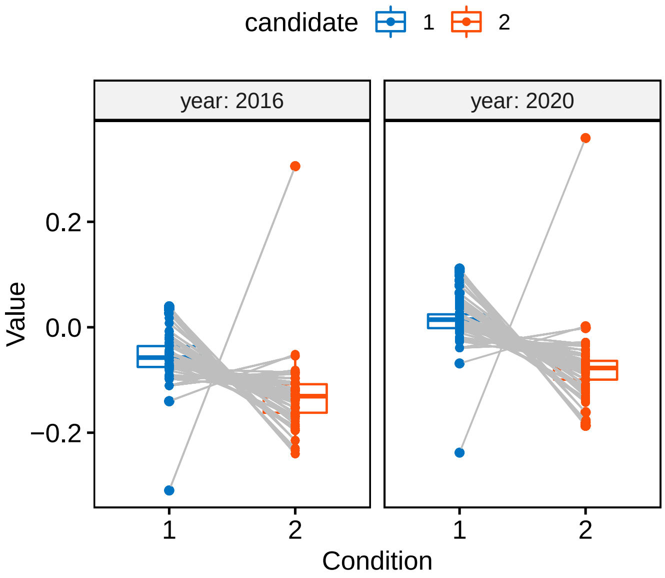
- Estimation results: Model (3)

Parameter	Estimate	Standard Error	t-statistic	Group
β_0	-0.049	0.010	-5.029	Fixed
β_1	-0.081	0.021	-3.905	Fixed
β_2	0.054	0.006	9.680	Fixed
β_3	-0.018	0.0071	-2.577	Fixed
σ	0.031	0.003	9.492	Pollster
$\sigma_{\rm d}$	0.060	0.007	8.469	State
$\sigma_{\rm r}$	0.093	0.010	9.227	State
ρ	-0.733	0.082	-8.895	State
τ	0.069	0.001	52.380	Residual

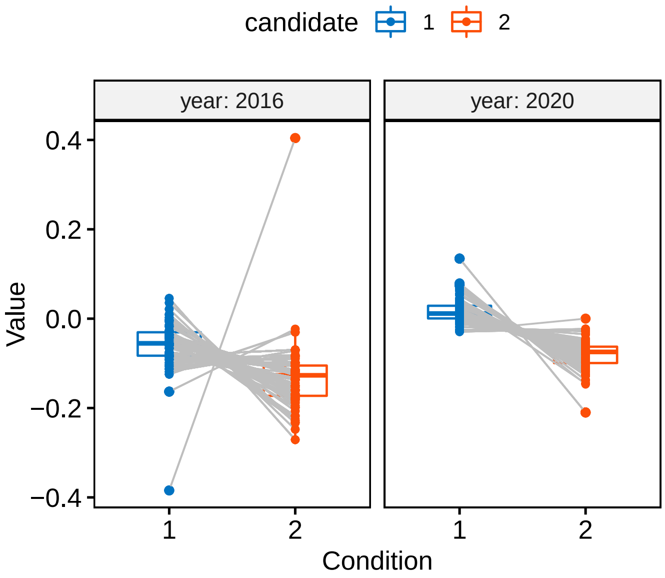
- Estimation results: Model (4)

Parameter	Estimate	Standard Error	t-statistic	Group
β_0	-0.055	0.010	-5.739	Fixed
β_1	-0.073	0.021	-3.573	Fixed
β_2	0.066	0.013	5.034	Fixed
β_3	-0.038	0.029	-1.302	Fixed
σ	0.032	0.003	9.935	Pollster
σ_{d}	0.058	0.006	10.358	Year:State
σ_{r}	0.091	0.007	12.659	Year:State
ρ	-0.798	0.054	-14.697	Year:State
τ	0.066	0.001	52.613	Residual

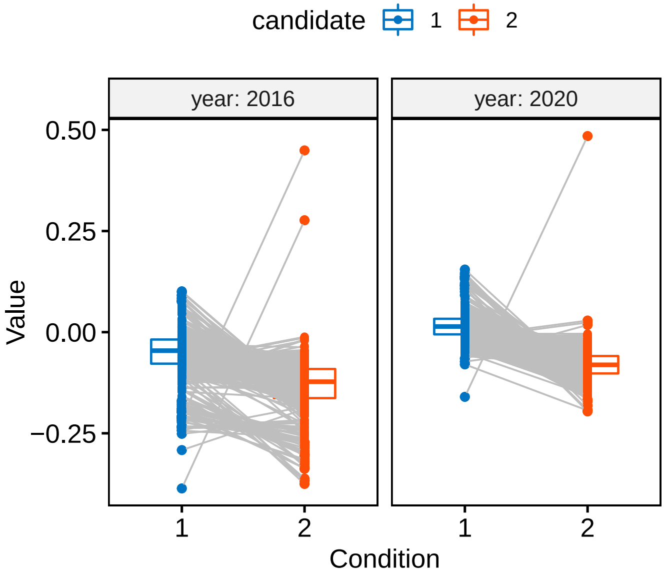
- Small area means: Model (1)



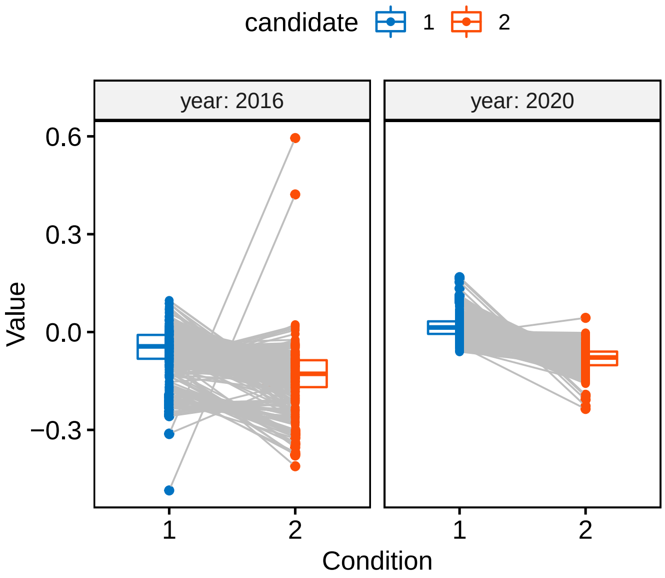
- Small area means: Model (2)



- Small area means: Model (3)

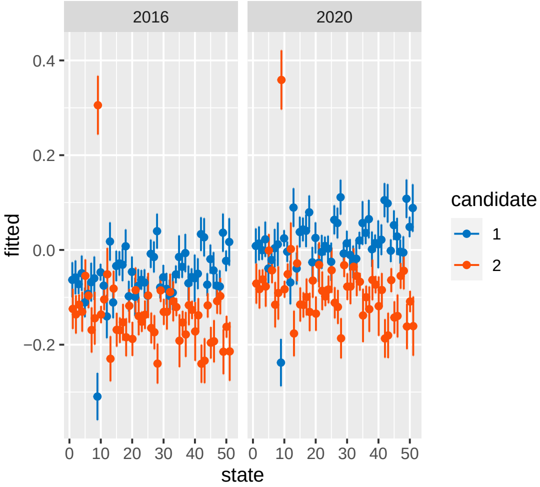


- Small area means: Model (4)

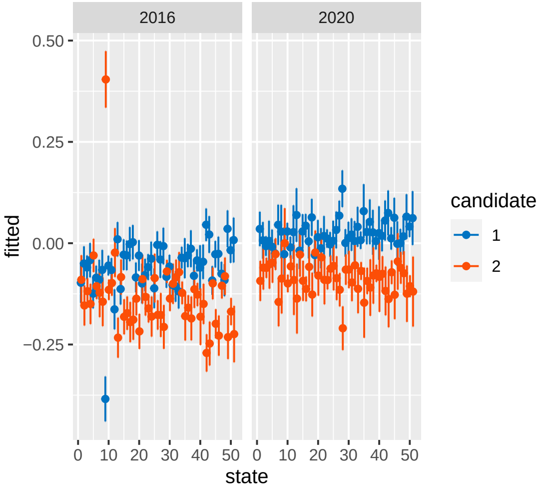


- Prediction intervals are constructed using the Sumca estimators (Jiang & Torabi 2020) of MSPE.

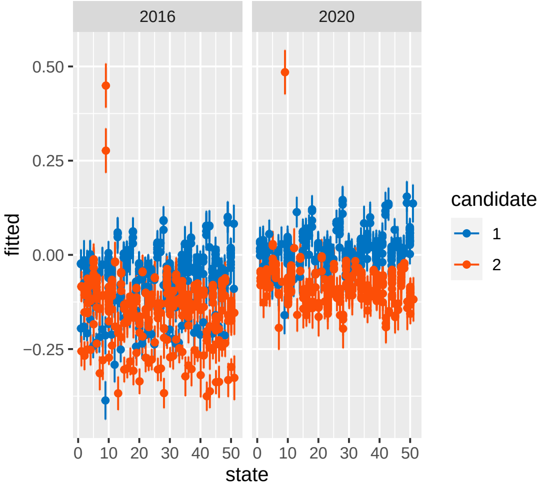
- Prediction intervals are constructed using the Sumca estimators (Jiang & Torabi 2020) of MSPE.
- Prediction intervals: Model (1)



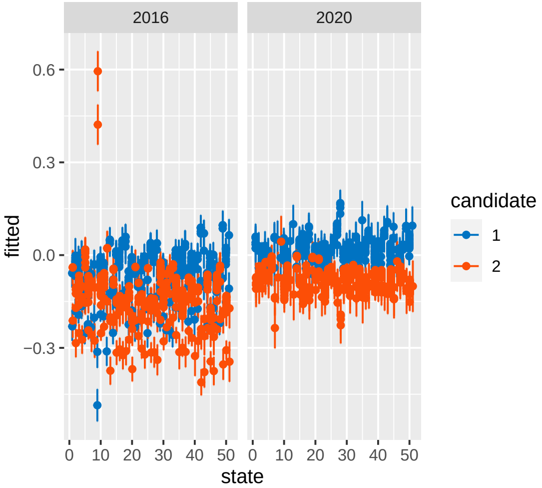
- Prediction intervals: Model (2)



- Prediction intervals: Model (3)



- Prediction intervals: Model (4)



3. Transfer Learning

- An interesting question is, given that there is strong similarity between the two elections in that both elections involved the same Republican candidate, could one have learned from one election to help improve prediction about another election?

3. Transfer Learning

- An interesting question is, given that there is strong similarity between the two elections in that both elections involved the same Republican candidate, could one have learned from one election to help improve prediction about another election?
- Practically, the elections, 2016 and 2020, are not at equal, or even similar, positions in terms of predicting one from the other, even if both predictions were possible.

3. Transfer Learning

- An interesting question is, given that there is strong similarity between the two elections in that both elections involved the same Republican candidate, could one have learned from one election to help improve prediction about another election?
- Practically, the elections, 2016 and 2020, are not at equal, or even similar, positions in terms of predicting one from the other, even if both predictions were possible.
- Reasons in terms of lesson-learning, or transfer learning.

3. Transfer Learning

- An interesting question is, given that there is strong similarity between the two elections in that both elections involved the same Republican candidate, could one have learned from one election to help improve prediction about another election?
- Practically, the elections, 2016 and 2020, are not at equal, or even similar, positions in terms of predicting one from the other, even if both predictions were possible.
- Reasons in terms of lesson-learning, or transfer learning.
- This lesson-learning prediction is simply our SAE method, now used to make prediction about the outcomes of another election involving the same Republican candidate.

- The "training data" consist of the election and poll results from one election (2016 or 2020), and the poll results from another election (2020 or 2016).

- The “training data” consist of the election and poll results from one election (2016 or 2020), and the poll results from another election (2020 or 2016).
- Because the data only involve one election year, Models (2) and (4) do not apply.

- The "training data" consist of the election and poll results from one election (2016 or 2020), and the poll results from another election (2020 or 2016).
- Because the data only involve one election year, Models (2) and (4) do not apply.
- Models 1 and 3 also need to be modified by dropping the year indicator and its interaction with the candidate indicator. The modified models are

$$\text{Model I}: \quad y_{ia} = \beta_0 + \beta_1 I_{\text{can},ia} + z_{ia} v_i + e_{ia},$$
$$\text{Model III}: \quad y_{ipc} = \beta_0 + \beta_1 I_{\text{can},ipc} + z_{ipc} v_i + u_p + e_{ipc}.$$

- First consider Model I. As noted, the small area means, $\theta_{ia} = \beta_0 + \beta_1 I_{\text{can},ia} + z_{ia} v_i$, do not depend on the pollster; therefore, they can be denoted by $\theta_{ik} = \beta_0 + \beta_1 1_{(k=2)} + v_i$, $i = 1, \ldots, 51$, $k = 1, 2$.

- First consider Model I. As noted, the small area means, $\theta_{ia} = \beta_0 + \beta_1 I_{\text{can},ia} + z_{ia} v_i$, do not depend on the pollster; therefore, they can be denoted by $\theta_{ik} = \beta_0 + \beta_1 1_{(k=2)} + v_i$, $i = 1, \dots, 51$, $k = 1, 2$.
- By fitting Model I with one year's data, we obtain the EBLUP,

$$\hat{\theta}_{ik} = \hat{\beta}_0 + \hat{\beta}_1 1_{(k=2)} + \hat{v}_i, \tag{5}$$

$i = 1, \dots, 51$, $k = 1, 2$.

- First consider Model I. As noted, the small area means, $\theta_{ia} = \beta_0 + \beta_1 I_{\text{can},ia} + z_{ia} v_i$, do not depend on the pollster; therefore, they can be denoted by $\theta_{ik} = \beta_0 + \beta_1 1_{(k=2)} + v_i$, $i = 1, \dots, 51$, $k = 1, 2$.
- By fitting Model I with one year's data, we obtain the EBLUP,

$$\hat{\theta}_{ik} = \hat{\beta}_0 + \hat{\beta}_1 1_{(k=2)} + \hat{v}_i, \tag{5}$$

  $i = 1, \dots, 51$, $k = 1, 2$.
- Then, for another year's election, we have, also under Model I,

$$\log(p^*_{ijk}) - \log(\pi^*_{ik}) \quad = \quad \theta^*_{ik} + e^*_{ijk}, \tag{6}$$

  where p^*_{ijk} is the poll result for state i, pollster j and candidate k, which is available for all i, j, k.

- We then average both sides of (6) over $j \in J_i$, where J_i is the subset of pollsters who ran the polls in state i, to get

$$\overline{\log(p^*_{i\cdot k})} - \log(\pi^*_{ik}) \quad = \quad \theta^*_{ik} + \bar{e}^*_{i\cdot k}, \tag{7}$$

where $\overline{\log(p^*_{i\cdot k})} = |J_i|^{-1} \sum_{j \in J_i} \log(p^*_{ijk})$, $\bar{e}^*_{i\cdot k} = |J_i|^{-1} \sum_{j \in J_i} e^*_{ijk}$, and $|J_i|$ denotes the cardinality of J_i.

- We then average both sides of (6) over $j \in J_i$, where J_i is the subset of pollsters who ran the polls in state i, to get

$$\overline{\log(p^*_{i\cdot k})} - \log(\pi^*_{ik}) \quad = \quad \theta^*_{ik} + \bar{e}^*_{i\cdot k}, \tag{7}$$

where $\overline{\log(p^*_{i\cdot k})} = |J_i|^{-1} \sum_{j \in J_i} \log(p^*_{ijk})$, $\bar{e}^*_{i\cdot k} = |J_i|^{-1} \sum_{j \in J_i} e^*_{ijk}$, and $|J_i|$ denotes the cardinality of J_i.

- If we replace the right side of (7) by the $\hat{\theta}_{ik}$ in (5), we obtain $\log(\pi^*_{ik}) \approx \overline{\log(p^*_{i\cdot k})} - \hat{\theta}_{ik}$, or, equivalently,

$$\pi^*_{ik} \approx \hat{\pi}^*_{ik} \equiv \exp\left\{\overline{\log(p^*_{i\cdot k})} - \hat{\theta}_{ik}\right\}, \quad i = 1, \dots, 51, \;\; k = 1, 2. \tag{8}$$

- Question: If (8) a better predictor than the "poll of polls", that is,

$$\bar{p}^*_{i\cdot k} = \frac{1}{|J_i|}\sum_{j\in J_i} p^*_{ijk}, \quad i = 1,\dots,51, \quad k = 1,2? \tag{9}$$

- Question: If (8) a better predictor than the "poll of polls", that is,

$$\bar{p}^*_{i\cdot k} = \frac{1}{|J_i|}\sum_{j\in J_i} p^*_{ijk}, \quad i = 1,\ldots,51, \quad k = 1,2? \tag{9}$$

- Comparison results: Model I

Training	Predicting	State	Actual	SAE	PoP	EC Votes
2016	2020	Arizona	D	R	D	11
		Florida	R	D	D	29
		Nebraska	R	R	D	5
		North Carolina	R	R	D	15
		National	D (306/232)	D (324/214)	D (355/183)	
2020	2016	Florida	R	R	D	29
		Michigan	R	D	D	16
		Nevada	D	R	D	6
		North Carolina	R	R	D	15
		Pennsylvania	R	R	D	20
		Wisconsin	R	D	D	10
		National	R (232/306)	R (252/286)	D (322/216)	

- Comparison results: Model III

Training	Predicting	State	Actual	SAE	PoP	EC Votes
2016	2020	Arizona	D	R	D	11
		Florida	R	D	D	29
		Nebraska	R	R	D	5
		North Carolina	R	R	D	15
		National	D (306/232)	D (324/214)	D (355/183)	
	2016	Florida	R	R	D	29
		Michigan	R	R	D	16
		Nevada	D	R	D	6
		North Carolina	R	R	D	15
		Pennsylvania	R	R	D	20
		Wisconsin	R	R	D	10
		National	R (232/306)	R (226/312)	D (322/216)	

4. Concluding Remarks

- Dishonesty in surveys has long been known as a challenging problem, whose impact on the survey results is difficult to correct (e.g., Chesney and Penny 2013).

4. Concluding Remarks

- Dishonesty in surveys has long been known as a challenging problem, whose impact on the survey results is difficult to correct (e.g., Chesney and Penny 2013).
- However, the results of our analyses show that, given that there is a future election involving the same candidate, whose prediction results were seriously off due to the dishonesty in surveys, the bias due to such dishonesty can be largely reduced using a relatively simple SAE method.

4. Concluding Remarks

- Dishonesty in surveys has long been known as a challenging problem, whose impact on the survey results is difficult to correct (e.g., Chesney and Penny 2013).
- However, the results of our analyses show that, given that there is a future election involving the same candidate, whose prediction results were seriously off due to the dishonesty in surveys, the bias due to such dishonesty can be largely reduced using a relatively simple SAE method.
- As a result, one can, at least, avoid making the same mistake twice.

4. Concluding Remarks

- Dishonesty in surveys has long been known as a challenging problem, whose impact on the survey results is difficult to correct (e.g., Chesney and Penny 2013).
- However, the results of our analyses show that, given that there is a future election involving the same candidate, whose prediction results were seriously off due to the dishonesty in surveys, the bias due to such dishonesty can be largely reduced using a relatively simple SAE method.
- As a result, one can, at least, avoid making the same mistake twice.
- As noted, compared to 2016, 2020 is a "small lesson" to learn.

- But even with such a small lesson learned via the SAE method, one could predict the correct outcome the next time, and therefore correct the bias due to the dishonest responses, at least in the final outcome.

- But even with such a small lesson learned via the SAE method, one could predict the correct outcome the next time, and therefore correct the bias due to the dishonest responses, at least in the final outcome.
- Note that the dishonesty is expected to have persisted in the next election due to the involvement of the same candidate.

- But even with such a small lesson learned via the SAE method, one could predict the correct outcome the next time, and therefore correct the bias due to the dishonest responses, at least in the final outcome.
- Note that the dishonesty is expected to have persisted in the next election due to the involvement of the same candidate.
- Imagine what one could do if a "big lesson" were learned, such as 2016.

- But even with such a small lesson learned via the SAE method, one could predict the correct outcome the next time, and therefore correct the bias due to the dishonest responses, at least in the final outcome.
- Note that the dishonesty is expected to have persisted in the next election due to the involvement of the same candidate.
- Imagine what one could do if a "big lesson" were learned, such as 2016.
- Clearly, the same lesson should be, could be, and have been learned, one way or the other.