Modeling Voting Dynamics in a Two-Party System: Person–To–Person Interactions and Media Effect

Tin Phan1, Gabriela Navas2,3, Caleb Ignace4, Christopher Kribs5, Carlos Castillo-Garsow6 and Victor Moreno7

1Department of Mathematical and Statistical Sciences, Arizona State University
2Department of Mathematics, University of Los Andes
3Department of Biomedical Engineering, University of Los Andes
4Department of Mathematics and Statistics, East Tennessee State University
5Department of Mathematics, University of Texas at Arlington
6Department of Mathematics, Eastern Washington University
7Simon A. Levin Mathematical, Computational and Modeling Sciences Center, Arizona State University

July 29, 2016

Abstract

The current 2016 US presidential primary election, characterized by many unexpected results, provides an interesting context to study how voters are influenced in deciding who to support. We address this question by developing a class of models driven either by the effect of mass media or by social interaction among voters and
members of the parties. The dynamics are modeled using four compartments with a transition matrix in describing the evolution of a discrete-time Markov chain. Each model is studied and fit to poll data from the 2012 and 2016 presidential elections using numerical methods. A comparison across elections indicates that the social influence of each group changes from one election to another, but response to media is similar in both cases.

1 Introduction

Driven by the unpredictable nature of the 2016 US presidential primary races within the Democratic and Republican parties, this work aims at giving a qualitative description and quantitative analysis of the underlying forces that lead the development of the 2012 and 2016 US presidential elections. The main focus of this research is to describe how person-to-person interaction and mass media consumption affect the voting dynamics of the election. For example, how does the interaction with a person who supports Donald Trump affect the voting behavior of others or, how well does media influence account for changes in voting preference of members of the Democratic and Republican parties?

Donald Trump has dominated all other Republicans in the race for the Republican nomination in 2016 without spending as much money as the other Republicans did [19]. We speculate that this is due in part to the high amount of publicity or media (television specifically) coverage Trump had over the other candidates [55, 40]. Not only has the GOP presidential candidate received the most media coverage for the least cost in the
current race, but he has also achieved the most primary votes of any Republican in history, exceeding Governor George W. Bush in 2000. He has also received the most votes against a GOP candidate in history [7].

A well-known mathematical approach for studying elections and voting behavior is the “Michigan-style approach” [17], where analysis on the behavior of voters is conducted using poll data. Additionally, factors that influence groups of voters to make decisions, such as social interaction, media consumption and the economy, are also studied extensively in past literature [17, 28, 15, 31]. Ultimately, these factors drive the voting behaviors behind any presidential election. Since person–to–person interactions and media coverage are two of the main ways in which the “campaign effect” of a candidate reach people [17], we will limit our focus to these two mechanisms and use poll data as our measure of voting behavior. In our two–party system, both candidates are competing for votes through the spread of an idea [3]: “it’s in your best interest to vote for me.” Thus, we can view voting as an epidemiological problem, in the sense that individuals can infect others to act as they do, i.e., voting is contagious [39]. This approach has been applied to many areas outside of epidemiology, such as fanatic behavior [11] and, directly relevant to our discussion, political parties [22, 50].

Person–to–person interactions can act as social pressure to push nonvoters to vote [25]. In fact, in campaign elections in general, the goal is not necessarily to change individuals’ minds, but to get them to vote [14, 16]. Moreover, social contacts are more likely to result in driving the participants’ opinions closer together than further apart, as observed
in [25]. This effect is implemented in numerous agent-based models (see [20], in which every individual randomly copies one of its neighbors’ opinions in each time step). In our model, we account for this effect and its opposite—interaction with an extremely polarized population can result in driving one subject’s opinion drastically away from the other one’s view. This implementation is motivated by the “divide-and-conquer” strategy adopted by Trump, who emphasizes the divisions inside his own party by appealing to the resentment of a defined group (“non-college educated working class receptive to economic nationalism”, according to Peter Trubowitz) in the electorate of the Republican party [54].

A model for the growth of political party membership through word-of-mouth recruitment was applied to the Labor, Scottish National, and Conservative UK parties and gave a narrative consistent with historical evidence in [22]. Yet, it couldn’t account for external political events, something we expect mass media to account for.

One of the suggested methods to model media influence on voter’s behaviors is TV exposure (see Zaller in [17]), which we use in combination with the TV mentions-by-candidate database in [55] and coverage studies in [40, 41]. Zaller proposes several models which determine voting preference of individuals at a given time as a function of media exposure at that particular time. We explore a similar approach, in which an average individual changes his/her voting position with respect to the current media output. Moreover, we consider distinctly the effect of positive and negative coverage about a candidate on the individuals.

Studies have shown that political system can be viewed in two parts: mass media first
influences “opinion leaders”, individuals who carry influential weight in their social circle. Therefore, their sphere of contacts acts closely as they do—this is known as “the two-step flow of communication” [24]. In fact, we could build a model that takes into account both effects, but it would make analyzing the individual effect more difficult.

Additionally, there has been much work on collective behavior: May sets up a framework for a simple majority decision [30]; Green and Gerber describes how to handle large populations in which single decisions are negligible, yet every individual enjoys (or suffers) the consequences of the collective decision [18]. Many researchers have developed agent-based voter models, such as Fernández-Garcia, who showed that voting dynamics in the US can be modeled as a miscopying of opinion in spatial environments [15]. Laver concluded that a strategy consisting on getting more voters indefinitely is less effective in the long run than one based on satisfying a “threshold” share of votes [25]. Halu analyzed the role of social networks in elections by considering each party as a network in which each agent is represented and capable of deciding whether to be active (i.e., vote) or inactive [20]. Merrill and Grofman used a spatial model that examined voters’ behavior using geometry. In their work, candidates and voters are represented as vectors which quantify their position about economy and politics and voting preference depending on the magnitude of the distance between the voter and candidate vectors [31].

Though these collective behavior, agent-based, and social network approaches of modeling voting behavior are out of the scope of this paper, we apply some of their ideas and results to address the following questions:
• Between mass media and person-to-person interaction, which mechanism better describes the dynamics of the 2012 and 2016 elections? Are the 2012 and 2016 presidential elections fundamentally different in the way people are influenced to vote?

• For both 2012 and 2016 elections, what changes in voting behavior (caused by media and interaction) have a stronger impact on the outcome of the election?

In the next section, we will give the formulation of our models and notations, state our assumptions for each model and introduce our methods for data fitting and sensitivity analysis. We will then display and analyze our results in their corresponding political context, which will be used to answer the research questions. Limitations of our approach and suggestions for possible future work are to be found in the closing section.

2 Methods

2.1 Model Formulation Framework

Even though the effect of a third party may be critical under some conditions (for example the presidential election in 2000, see [50]), we focus on fluctuations between the two major parties. Thus, we consider a two-party-system for our model, e.g. a Democratic-Republican system, and a population in which every individual belongs to one party.

The considered population is divided into four classes:

• \( V_1(t) \) - Voters for Rep. candidate.
• \( M_1(t) \) - Non-voting Rep. member.
• $V_2(t)$ - Voters for Dem. candidate.  
• $M_2(t)$ - Non-voting Dem. member.

$V_1(t)$ and $V_2(t)$ represent the total percentage of people within the two parties who report themselves to be voting for the Republican or Democratic candidate at time $t$, respectively. The classification of $V_1$ and $V_2$ is specific to the candidate of each party. However, we do not directly examine the individuality of each candidate, e.g., their behaviors, credibility or ideology. Instead, we use our models and empirical results to ascertain the most important movement(s) within the population in each election. $M_1(t)$ and $M_2(t)$ are people who self-identified with the Republic and Democratic party but do not have intentions of voting for either party at time $t$, respectively. For the rest of the paper, we will refer to these categories as $V_1$, $V_2$, $M_1$, and $M_2$ for the purpose of reducing clutter in equations, time being implicit. The total population, $N$, only consists of people from the two parties, so $N = V_1 + V_2 + M_1 + M_2$ and $\frac{dN}{dt} = 0$. We use an epidemiological approach to characterize the different interactions and population transition between each group. Similar applications have been done recently in [53] and [50]. We treat influence similar to infectivity in the sense that people in some groups can drive other people to change class. However, influence can both attract and repulse others, whereas infectious individuals usually attract people from different classes to their own group (or a similar one).

We consider six types of possible transitions in one time-step, as in Figure 1:

• A non-voting member becomes a voting member of the same party, $M_i \rightarrow V_i$.

• A non-voting member becomes a voting member of the other party, $M_i \rightarrow V_j$. 

7
• A non-voting members become a non-voting members of the other party, $M_i \rightarrow M_j$.

• A voter becomes a non-voting a member of the same party, $V_i \rightarrow M_i$.

• A voter switches side to vote for the other candidate, $V_i \rightarrow V_j$.

• An individual stays its group, $V_i \rightarrow V_i$ or $M_i \rightarrow M_i$.

Notice that there’s no direct transition a between voting class and the opposite member class. This comes from the assumption that, if committed voters change their candidate preference, they will become committed voters for the other candidate ($V$), not merely sympathetic ($M$). This is supported by extending the conclusion that once a person votes, it increases the chance of that person voting in the future to once a person intents to vote, it is likely that that person will keep his/her intention in the future [18]. however, committed voters who become disillusioned are assumed to retain their general party preference.

![Graphical representation of movement within the system due to either media or interaction mechanisms](image)

Figure 1: **Graphical representation of movement within the system due to either media or interaction mechanisms**

We set our model as a discrete-time Markov chain, $\pi(t + 1) = \pi(t)T$, with $\pi =$
being the state vector. The transition matrix $T$ is defined as follows:

$$
T = \begin{pmatrix}
P_{V_1}^{V_1} & P_{M_1}^{V_1} & 0 & P_{V_2}^{V_1} \\
P_{M_1}^{M_1} & P_{M_1}^{M_1} & P_{M_2}^{M_1} & P_{V_2}^{M_1} \\
P_{M_1}^{M_2} & P_{M_2}^{M_1} & P_{M_2}^{M_2} & P_{V_2}^{M_2} \\
0 & P_{V_2}^{V_2} & P_{V_2}^{M_1} & P_{V_2}^{M_2}
\end{pmatrix}
$$

(1)

where $P_y^x$ is the probability of an individual from the superscript group, $y$, moving to the subscript group, $x$.

### 2.1.1 Person–to–Person Interaction Model

Graphical representations of the possible movements within the system and the interactions under consideration are in Figure 1 and 2, respectively.

The following assumptions are taken within the person–to–person interaction model:

- All movements are influenced by person–to–person contact.

- All individuals are influential but with different degree of influence. In particular, people from the $V$ groups exert influence with more extreme effect than people from the $M$ groups. This does not assume that $V$ groups are more capable of changing the political affiliation of someone else. Instead, it indicates the possible extreme results that arise from interacting with $V$, e.g. $V_1 \rightarrow V_2$.

- The population is homogeneously mixed, e.g., an individual is connected to all others and is equally likely to make contact with any other individual. This means the chance of coming into contact with an individual in the $V_1$ population is $V_1/N$, with
\[ N = V_1 + M_1 + M_2 + V_2. \] This assumption is partially supported by the rise of social networks, in particular Facebook, which facilitate connection between people with all kinds of political opinions.

- It is possible for non-voting members to be influenced to vote and for voting members to stop voting.

Figure 2: **Person-to-person interaction and resulting effects**: The dashed arrows represent the interaction and the solid lines represent the result(s) of the interaction.
To illustrate an interaction (which exists in Figure 2), take $V_1 \rightarrow M_1 \rightarrow V_2$. This means "$V_1$ interacts with $M_1$ causing $M_1$ to change to $V_2$." The probability of the transition is then $P(V_2|M_1,V_1)$, i.e., the probability that an individual is $V_2$ at time $(t + 1)$ given previously being $M_1$ at and having an interaction with $V_1$ at time $t$. We can now update the definition of the probability of moving from group $y$ to group $x$ to account for all interactions:

$$P_y^x = \sum_{k \in \pi} \left( P(y|x,k) \frac{k}{N} \right),$$

where $\pi = (V_1, M_1, M_2, V_2)$.

For convenience, the following notations are defined for the parameters in the context of one time step:

- $\alpha^y_x$: probability of an individual from the superscript group $y$ not moving after an interaction with an individual from the subscript group $x$—a failure to influence, where $x$ is the source of influence.

- $\beta^y_x$: probability of an individual from the superscript group $y$ moving to a group of the same party, given that the interaction with an individual from the subscript group $x$ is a success, from $x$’s perspective in a time step.

- $\gamma^y_x$: probability of an individual from the superscript group moving to the $M$-group of the other party (the party which $y$ belongs to), given that the interaction with the subscript group is a success.
All probabilities $P_{vy}$ of the interaction transition matrix are as follows:

\[
P_{V_1} = \alpha_{V_1} \frac{V_1}{N} + \alpha_{V_2} \frac{V_2}{N} + \alpha_{M_1} \frac{M_1}{N} + \frac{M_2}{N}
\]

\[
P_{M_1} = (1 - \alpha_{M_1}) \beta_{V_1} \frac{V_1}{N} + (1 - \alpha_{M_1}) \beta_{M_1} \frac{M_1}{N}
\]

\[
P_{M_2} = (1 - \alpha_{M_2}) (1 - \beta_{V_1}) (1 - \gamma_{M_2}) \frac{V_1}{N} + (1 - \alpha_{M_2}) (1 - \beta_{M_2}) \frac{M_1}{N}
\]

\[
P_{V_2} = (1 - \alpha_{V_1}) (1 - \beta_{V_2}) \frac{V_1}{N}
\]

\[
P_{M_1} = (1 - \alpha_{V_1}) \beta_{V_2} \frac{V_2}{N} + (1 - \alpha_{V_2}) \gamma_{M_2} \frac{M_1}{N}
\]

\[
P_{M_2} = (1 - \alpha_{V_1}) (1 - \beta_{V_2}) (1 - \gamma_{M_2}) \frac{V_2}{N} + (1 - \alpha_{V_2}) \frac{M_2}{N}
\]

\[
P_{V_2} = (1 - \alpha_{V_1}) \frac{V_1}{N} + \alpha_{V_2} \frac{V_2}{N} + \alpha_{M_2} \frac{M_1}{N} + \frac{M_2}{N}
\]

\[
P_{V_2} = (1 - \alpha_{V_1}) (1 - \beta_{V_2}) \frac{V_2}{N}
\]

\[
P_{M_2} = (1 - \alpha_{M_1}) (1 - \beta_{V_1}) \frac{V_1}{N} + (1 - \alpha_{M_2}) (1 - \beta_{V_2}) (1 - \gamma_{M_1}) \frac{V_2}{N}
\]

\[
P_{V_2} = (1 - \alpha_{M_1}) \beta_{M_2} \frac{V_1}{N} + (1 - \alpha_{M_2}) \frac{M_2}{N}
\]

\[
P_{V_2} = \alpha_{V_1} \frac{V_1}{N} + \alpha_{V_2} \frac{V_2}{N} + \alpha_{M_2} \frac{M_2}{N} + \frac{M_1}{N}
\]

### 2.1.2 Media Model

This model assumes that members and voters switch class motivated only by the amount of favorable or unfavorable media coverage of either candidate. TV mentions of different candidates by source are to be found in [55] and reports from Pew Research Center [41]
have quantified the amount of favorable or negative coverage of candidates over time.

Based on these data, we can define the following:

- $f_i(t)$ is the favorable coverage (e.g. TV mentions) for candidate $i$;
- $n_i(t)$ is the negative coverage for candidate $i$;
- $T(t) = f_1 + n_1 + f_2 + n_2$ is the total considered coverage at time $t$;

and

$$F_i(t) \equiv \frac{f_i(t)}{T(t)}, \quad N_i(t) \equiv \frac{n_i(t)}{T(t)}.$$

We assume that enough negative coverage of one candidate alone can make voters decay—loss of preference—to members ($V_i \rightarrow M_i$) or switch to vote for the opponent.

Members can be recruited by favorable coverage of the candidate or by negative coverage of the opponent.

The parameters are defined as follow:

- $\sigma_i$: effectiveness of $N_i$ on the movement from $V_i$ to $V_j$
- $\delta_i$: effectiveness of $N_i$ on the movement from $V_i$ to $M_i$
- $\gamma_i$: effectiveness of $N_j + F_i$ on the movement from $M_i$ to $V_i$.
- $\lambda_i$: effectiveness of $N_i + F_j$ on the movement from $M_i$ to $V_j$.
- $\kappa_i$: effectiveness of $N_i + F_j$ on the movement from $M_i$ to $M_j$. 
Figure 3: Media model - a map of movements within the system

The equations for passing from one class to the other are presented in Figure 3. Here $\delta_i$, $\gamma_i$, $\lambda_i$, $\sigma_i$, $\kappa_i$ are constants which play the role of amplification or reduction factors. This means that, for example, the amount of negative coverage about candidate $i$ drives a voter in $V_i$ towards $M_i$, but the effectiveness of the coverage depends on $\delta_i$. If $\delta_i = 0$, then individuals won’t go from $V_i$ to $M_i$, no matter how much negative coverage the candidate $i$ has. Otherwise, the transition probability with respect to the amount of coverage will be linear. In a sense, they represent how “sensitive” a group is to the influence of the media.

In this model, the transition probabilities are dependent on the media coverage over
Now let the parameter vector be defined as follows:

\[
\vec{\rho} = [\sigma_1 \sigma_2 \delta_1 \delta_2 \gamma_1 \gamma_2 \lambda_1 \lambda_2 \kappa_1 \kappa_2]
\]

Then two constraints must be satisfied:

- \( \bar{0} \leq \vec{\rho} \leq \bar{1} \)

- Each row sum in the transition matrix is equal to 1. This restriction can be stated by defining the matrix

\[
A = \begin{bmatrix}
N_1 & 0 & N_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & N_2 & 0 & N_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & N_2 + F_1 & 0 & N_1 + F_2 & 0 & N_1 + F_2 & 0 \\
0 & 0 & 0 & 0 & N_1 + F_2 & 0 & N_1 + F_1 & 0 & N_2 + F_1 & 0 \\
\end{bmatrix}
\]

15
and setting $0 \leq A_i^{j'} \leq 1$ in the parameter estimation process, so that the sum of the non-diagonal elements in each row of the transition matrix is between 0 and 1.

2.1.3 Data collection

We use two different data sets (polling data) that contain information on voting preferences from the people who self-identify as either Democrat or Republican during the 2012 and 2016 presidential elections for parameter fitting. We further cross reference the credibility of each data source with the rating published by Nate Silver [52]. All polls have a rating of $A-$ or higher, with the exception of Politico/George Washington University with a $B$ rating.

Specifically, the first data set is compiled of weekly\(^1\) data sets from multiple sources spanning 14 weeks from the beginning of August of 2012 to right before the election day, November 6, 2012, from [1, 8, 9, 10, 12, 42, 43, 44, 45, 46, 47, 48, 49]. All polls are nationwide polls with the exception of one statewide poll, which still fits the trend of the data. The second set is collected through a single source. In this case, the data are taken weekly\(^2\) between May 18 to July 12 of 2016 from [32, 33, 34, 35, 35, 36, 37, 38].

We also collect data on the number of times a candidate is mentioned on TV and headlines either favorably or unfavorably. For the 2012 election, data is taken from the studies in [41]. For 2016, the period from May to June is covered by the report in [40]. Data on the month of July is taken from [55].

2.2 Computational Methods

In order to determine the driving mechanisms behind an election, we first need to find out which parameter has the greatest impact under each assumption. To do this, we use two different methods of parameter fitting and compare the results to one another.

\(^1\)We use Monday as a reference. If the data set is taken within that week, then we categorize it as data for that week.

\(^2\)Wednesday is used as the reference similarly.
2.2.1 Parameter estimation

For the media model, we assume constant media coverage between consecutive data points. We then update the parameters using the favorable and unfavorable coverage data. The model is fitted through a minimization of the least squared error function using a minimizing MATLAB built-in function ("fmincon") that allows for constrains on the minimizing parameters [29] to find the proportional constants that scale the effect of the media to changes in voting behavior. Note that a global minimum is not guaranteed.

For the person-to-person interaction model, we first apply the same scheme as above. However, due to the complexity of the transition matrix (state-dependency and non-linearity), many local error minima may occur. This could potentially give bad parameter estimates. Thus, we run the function multiple times with random initial guesses to obtain multiple sets of best fit parameters (with the same fitting error) for each data set and choose one that is representative of each set. Each set of parameters varies only slightly from one another. This suggests that we either obtain a global minimum or stuck in a local minimum. Some of the variations in each set of parameters are artificial in the sense that they do not affect the fitting. More precisely, since the conditional probabilities $\beta$ and $\gamma$ only matter if the corresponding $\alpha$ is not 1, it means that some variations occur in $\beta$ and $\gamma$, when the corresponding $\alpha$ is 1.

2.2.2 Sensitivity Analysis

To reach our goal of finding the driving force behind each election, we carry out sensitivity analysis for both systems [2]. We define the outcome of the election as

$$V = V_1(t_E) - V_2(t_E)$$

where $t_E$ is the week of the election (which we take to be the last time-step of each simulation). This quantity is negative if the Democrat candidate wins, and positive if the Republican candidate does. The magnitude represents the difference between the percent-
age of votes each candidate got.

Interaction model

For the interaction model, we calculate the closed forms of the sensitivity of $V \equiv V_1(t + 1) - V_2(t + 1)$ with respect to each parameter between two consecutive time points. Recall that $V_1, V_2, M_1, M_2$ are taken to be evaluated at time $t$.

\begin{align*}
\frac{a_1}{V} \frac{\partial V}{\partial a_1} &= \frac{V^2}{V} \\
\frac{a_2}{V} \frac{\partial V}{\partial a_2} &= \frac{V_1 M_1 (1 - 2b_1) a_2}{V} \\
\frac{a_3}{V} \frac{\partial V}{\partial a_3} &= \frac{V_1 M_2 (2b_3 + c_1 - b_2 c_1 - 1) a_3}{V} \\
\frac{a_4}{V} \frac{\partial V}{\partial a_4} &= \frac{V_2 V_1 (b_3 - 2) a_4}{V} \\
\frac{a_5}{V} \frac{\partial V}{\partial a_5} &= \frac{V_1 M_1 a_5}{V} \\
\frac{a_6}{V} \frac{\partial V}{\partial a_6} &= 0 \\
\frac{a_7}{V} \frac{\partial V}{\partial a_7} &= \frac{V_1 V_2 (2 - b_4) a_7}{V} \\
\frac{a_8}{V} \frac{\partial V}{\partial a_8} &= \frac{V_2 M_1 (b_2 c_2 - 2b_5 - c_2 - 1) a_8}{V} \\
\frac{a_9}{V} \frac{\partial V}{\partial a_9} &= \frac{V_2 M_2 (2b_6 - 1) a_9}{V} \\
\frac{a_{10}}{V} \frac{\partial V}{\partial a_{10}} &= -\frac{V_2^2 a_{10}}{V} \\
\frac{a_{11}}{V} \frac{\partial V}{\partial a_{11}} &= 0 \\
\frac{a_{12}}{V} \frac{\partial V}{\partial a_{12}} &= -\frac{V_2 M_2 a_{12}}{V} \\
\frac{b_1}{V} \frac{\partial V}{\partial b_1} &= \frac{2 V_1 M_1 (1 - a_2) b_1}{V} \\
\frac{b_2}{V} \frac{\partial V}{\partial b_2} &= -\frac{M_2 V_1 (1 - a_3) (2 - c_1) b_2}{V} \\
\frac{b_3}{V} \frac{\partial V}{\partial b_3} &= \frac{V_2 V_1 (a_4 - 1) b_3}{V} \\
\frac{b_4}{V} \frac{\partial V}{\partial b_4} &= \frac{V_1 V_2 (1 - a_7) b_4}{V} \\
\frac{b_5}{V} \frac{\partial V}{\partial b_5} &= \frac{M_1 V_2 (1 - a_8) (2 - c_2) b_5}{V} \\
\frac{b_6}{V} \frac{\partial V}{\partial b_6} &= \frac{2 M_2 V_2 (a_9 - 1) b_6}{V} \\
\frac{c_1}{V} \frac{\partial V}{\partial c_1} &= -\frac{M_2 V_1 (1 - a_3) (1 - b_2) c_1}{V} \\
\frac{c_2}{V} \frac{\partial V}{\partial c_2} &= \frac{M_1 V_2 (1 - a_8) (1 - b_5) c_2}{V}
\end{align*}

so that the normalized sensitivity of $V(t_E)$ is obtained by evaluating $V$ in the last time-step.

Media model

To find the sensitivity values in the media model, we vary each parameter by 1% while
fixing all of the others and use the model to recalculate $V$. Let the percent change in $V$ be

$$\%V = \frac{V_{est} - V_{rec}}{V_{est}},$$

where $V_{est}$ is the value of $V$ with the original estimated set of parameters and $V_{rec}$ is the value obtained after modifying the parameter. Sensitivity value will then be the relation between percent change in $V$ and percent change in the parameter [6].

3 Results

3.1 Interaction Models

3.1.1 2012 Election

For the interaction model, the best fit curves of the 2012 poll data is shown in Figure 4 with SSE of 0.0149. The parameter estimates are shown in Table 1. Interestingly, there are many zero and one probabilities. These parameter values define a set of implications on the result of interactions between two individuals in the system. These are explained in full below and collectively visualized in Figure 14. Note that for the following, the usage of *always* and *never* are approximations by our model. The following conclusions only address the result of one encounter between two individuals in a single time-step.

Since $\alpha_{V_1}^{V_1}, \alpha_{V_2}^{V_1} > 0.5$, $\alpha_{M_1}^{V_1} = 0$, and $\alpha_{M_2}^{V_1} = 1$, a Romney voter (a person who declares to be voting for Romney, despite the party he/she belongs to):

- *usually* remains a supporter of Romney following an interaction with either another Romney voter or Obama voter.
- *never* remains a supporter of Romney following an interaction with a non-voting Republican member (that is a person who self-identify with the Republican party but do not wish to support Romney)
- *always* remains a supporter of Romney following an interaction with a Democratic member.
Since $\alpha_{V_2} = 1$, $\alpha_{V_1} > 0.5$, $\alpha_{M_1} = 0$, and $\alpha_{M_2} < 0.5$, an Obama voter

- *always* remains a supporter of Obama following an interaction with another Obama voter.
- *usually* remains a supporter of Obama following an interaction with a Romney voter.
- *usually doesn’t* remain a supporter of Obama following an interaction with a Democratic member.
- *never* remains a supporter of Obama following an interaction with a Republican member.

Since $\alpha_{M_1} < 0.5$ and $\alpha_{M_2} = 0$, a Republican member

- *usually doesn’t* stay uncommitted after an interaction with a Romney voter.
- *never* stays uncommitted following an interaction with a Obama voter.

Lastly, $\alpha_{V_1} = 0$ and $\alpha_{V_2} = 1$, so a non-voting Democratic member

- *never* stays uncommitted following an interaction with a Romney voter.
- *always* stays uncommitted following an interaction with a Obama voter.

Substituting the estimated parameters in the transition matrix, we obtain

$$T = \begin{pmatrix}
\frac{0.99V_1+M_2+0.96V_2}{N} & \frac{0.01V_1+M_1}{N} & 0 & \frac{0.04V_2}{N} \\
0 & \frac{0.02V_1+M_1+M_2}{N} & 0 & \frac{0.98V_1+V_2}{N} \\
0 & \frac{V_1+M_1}{N} & \frac{M_2+V_2}{N} & 0 \\
\frac{0.09V_1}{N} & 0 & \frac{0.77M_2}{N} & \frac{0.91V_1+M_1+0.23M_2+V_2}{N}
\end{pmatrix}$$

This gives some important results about the fitted model:

- Republican members cannot directly (in one time step) become Democrat members.
Figure 4: **Interaction model and best fit curves for the 2012 election:** The plot on the top shows the modeled voting classes, $V_1$ and $V_2$, and the bottom plot presents the modeled member classes, $M_1$ and $M_2$. 
Table 1: Parameter estimates and sensitivities for the interaction model applied to the 2012 election

- There is no direct recruitment of party members to party voters within either party.
- Democrat members can only change to become Republican members.
- Republican voters usually drive Republican members away to vote for the other party.
- When voters from different parties interact, the Democrat voter is around twice as likely to become a Republican voter than the Republican is to switch candidates.
- Democrat voters can decay to members by interaction with Democrat members, while Republican voters can decay by interaction with both Republican members and voters.

The sensitivity analysis, see Table 1 and Figure 5, reveals that $\alpha_{V_1}$ and $\alpha_{V_1}$, which
Figure 5: Interaction model and bar plot for the contrast of parameter sensitivity to the parameter value for the 2012 election: Y-axis is in log-scale.

represent the probability that a voter interacts with a voter of the opposite party and doesn’t change class, have the most impact on the outcome of the election. Other than these, the parameters with higher sensitivity are $\alpha_{V_1}^V$ and $\alpha_{V_2}^V$ (i.e. probability of voters staying the same given that they interact within their own class). Note that the interactions that most affect the change in the final vote are meetings between voters of each party; the election is strongly dependent on how likely voters are to fail to change each other’s minds.

3.1.2 2016 Election

The best fit curve for the 2016 poll data is shown in Figure 6 with SSE of 0.0037. We have the estimation for the parameters and the sensitivity in the last time step recorded in Table 2. Like the 2012 election parameters values, the 2016 election parameters values have meanings that are explained below and are visualized in Figure 14.
Since $\alpha_{V_1}^{V_1} = 1$, $\alpha_{V_1}^{V_2} < 0.5$, $\alpha_{M_1}^{V_1} = 1$, and $\alpha_{M_2}^{V_1} = 1$, a voter for Trump

- always remains a supporter of Trump due to interaction with either another Trump voter or Republican or Democratic member.

- usually doesn’t remain a supporter of Trump after an interaction with a Clinton voter.

Since $\alpha_{V_2}^{V_2} = 1$, $\alpha_{V_2}^{V_1} < 0.5$, $\alpha_{M_2}^{V_2} = 0$, and $\alpha_{M_2}^{V_2} = 1$, a voter for Clinton

- always remains a supporter of Clinton after an interaction with either another Clinton voter or non-voting Democratic member.

- usually doesn’t remain a supporter of Clinton after an interaction with a Trump voter.

- never remains a supporter of Clinton after an interaction with a Republican member.

Since $\alpha_{M_1}^{M_1} = 0$ and $\alpha_{M_1}^{M_2} = 0$, a Republican member

- never stays uncommitted after an interaction with a Trump voter.

- never stays uncommitted after an interaction with a Clinton voter.

Lastly, $\alpha_{V_1}^{M_2} = 1$ and $\alpha_{V_2}^{M_2} = 0$, so a Democratic member

- always stays uncommitted after an interaction with a Trump voter.

- never stays uncommitted after an interaction with a Clinton voter.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimation</th>
<th>Sensitivity</th>
<th>Parameter</th>
<th>Estimation</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{V_1}$</td>
<td>1</td>
<td>0.235</td>
<td>$\alpha_{V_2}$</td>
<td>0.4670</td>
<td>0.188</td>
</tr>
<tr>
<td>$\alpha_{M_1}$</td>
<td>0</td>
<td>0</td>
<td>$\beta_{M_1}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\alpha_{M_2}$</td>
<td>1</td>
<td>0.013</td>
<td>$\alpha_{M_2}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\alpha_{V_2}$</td>
<td>0.3441</td>
<td>-0.159</td>
<td>$\alpha_{V_2}$</td>
<td>1</td>
<td>-0.417</td>
</tr>
<tr>
<td>$\alpha_{V_1}$</td>
<td>1</td>
<td>0.070</td>
<td>$\alpha_{M_1}$</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$\alpha_{M_2}$</td>
<td>0</td>
<td>0</td>
<td>$\alpha_{M_2}$</td>
<td>1</td>
<td>-0.093</td>
</tr>
<tr>
<td>$\beta_{V_1}$</td>
<td>1</td>
<td>0.140</td>
<td>$\beta_{V_2}$</td>
<td>0.7123</td>
<td>0.119</td>
</tr>
<tr>
<td>$\beta_{V_2}$</td>
<td>0.4625</td>
<td>0</td>
<td>$\beta_{V_2}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{V_1}$</td>
<td>0.5228</td>
<td>-0.107</td>
<td>$\beta_{M_2}$</td>
<td>1</td>
<td>-0.187</td>
</tr>
<tr>
<td>$\gamma_{M_2}$</td>
<td>0.4852</td>
<td>0</td>
<td>$\gamma_{M_1}$</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2: Parameter estimates and sensitivities for the interaction model applied to the 2016 election

The transition matrix we arrive at with the parameter values is

$$ T = \begin{pmatrix}
\frac{V_1 + M_1 + M_2 + 0.47V_2}{N} & \frac{0.38V_2}{N} & 0 & \frac{0.15V_2}{N} \\
\frac{V_1}{N} & \frac{M_1 + M_2}{N} & 0 & \frac{V_2}{N} \\
0 & \frac{M_1}{N} & \frac{V_1 + M_2}{N} & \frac{V_2}{N} \\
\frac{0.32V_1}{N} & 0 & \frac{0.34V_1}{N} & \frac{0.34V_1 + M_1 + M_2 + V_2}{N}
\end{pmatrix} $$

Again, we can read some information from the transition matrix, written in contrast to the 2012 election transition matrix (section 3.1):

- Republican members cannot directly (in one time step) become Democrat members.
- There is direct recruitment of party members to party voters within either party.
- If voters from different parties interact, the Democrat is still nearly twice as likely as the Republican to switch candidates.
Figure 6: **Interaction model and best fit curves for the 2016 election**: The plot on the top shows the modeled voting classes, $V_1$ and $V_2$, and the bottom plot presents the modeled member classes, $M_1$ and $M_2$.

- Both voter groups only decay to their party’s member groups by interaction with the other voter group.

- Every transition between different groups is solely caused by a voter class, except for the case in which Republican members influence Democratic members to become Republican members.

Note also that Republican voters no longer influence voters or members in their own party to stop supporting the candidate or leave the party. Yet, Republican voters no longer influence Democratic members to become Republican members.

The sensitivity analysis, see Table 2 and Figure 7, reveals the order of the five parameters to which the outcome of the final poll is most sensitive: $\alpha_{V_1}^{V_2}$, (the probabilities
that $V_1$ stays $V_1$ given an interaction with $V_2$, $\alpha_{V_1}^{V_2}$, $\alpha_{V_1}^{V_1}$, and $\beta_{V_1}^{V_2}$. As in the 2012 election interaction model, we find that the most sensitive parameters are those associated with the likelihood of voters failing to changing other voters’ minds.

### 3.2 Media Models

#### 3.2.1 2012 Election

The parameter estimation for the media model produces the best-fit line in figure 8. Table 3 presents the results. The parameter with the highest fitted value is $\gamma_2$, which means that the media effect has more influence in the Democrat members’ recruitment (i.e. the transition from $M_2$ to $V_2$) than in other movements. The parameters with lower values reveal the transitions in which the media effect has a lower impact. In this case, there are four parameters with values of order $10^{-5}$ or $10^{-6}$, namely $\gamma_1, \lambda_2, \delta_1$ and $\sigma_1$. The
Figure 8: **Media model and best fit curves for the 2012 election**: The plot on the top shows the modeled voting classes, $V_1$ and $V_2$; the middle plot presents the modeled member classes, $M_1$ and $M_2$; the bottom plot displays the amount of media coverage at a given time. For each candidate, we present the percentage of negative coverage affecting him and the added percentage of his own negative coverage plus the positive coverage for the competitor.
corresponding transitions with low media susceptibility are:

- Recruitment ($M_1$ to $V_1$) inside the Republican party;
- Recruitment ($M_2$ to $V_1$) outside the Republican party;
- Decaying ($V_1$ to $M_1$) inside the Republican party;
- Switching ($V_1$ to $V_2$) from voting Republican to voting Democrat.

We wish to know which parameters have more impact in determining the final result of the election. We vary each estimated parameter individually by a few percent and calculate the effect it has on the outcome of the election, namely $V_1 - V_2$. The calculated sensitivity indexes are presented in Table 3 and Figure 9. Parameters driving transitions which decrease the number of Republican voters and/or increase the number of Democratic voters naturally have negative sensitivity indices for $V$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Sensitivity</th>
<th>Parameter</th>
<th>Estimate</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_1$</td>
<td>8.90e-6</td>
<td>-0.0011</td>
<td>$\delta_2$</td>
<td>4.68e-2</td>
<td>0.8290</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>5.35e-5</td>
<td>0.0004</td>
<td>$\gamma_2$</td>
<td>8.47e-1</td>
<td>-0.9187</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>2.35e-4</td>
<td>-0.0003</td>
<td>$\lambda_2$</td>
<td>6.13e-5</td>
<td>0.0006</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>9.69e-6</td>
<td>-0.0016</td>
<td>$\sigma_2$</td>
<td>1.35e-2</td>
<td>2.2896</td>
</tr>
<tr>
<td>$\kappa_1$</td>
<td>6.85e-1</td>
<td>-0.4012</td>
<td>$\kappa_2$</td>
<td>4.25e-1</td>
<td>0.3321</td>
</tr>
</tbody>
</table>

Table 3: Parameter estimates and sensitivities for the media model applied to the 2012 election. The SSE for the estimation is 0.0137.
Figure 9: Media model and bar plot for the contrast of parameter sensitivity to the parameter value for the 2012 election: Y-axis is in log-scale.

First note that the five parameters with lower values are actually the ones to which the outcome is less sensitive. However, the largest parameter, $\gamma_2$, does not produce the highest sensitivity value: $\sigma_2$, which is not among the highest values, does. This last fact suggests that the outcome of the election was most affected by how likely negative coverage for the Democratic candidate is to drive Democrat voters to vote Republican. The parameter $\gamma_2$ being the largest means that recruitment in the Democrat party is the transition most affected by media. The next values in magnitude are the $\kappa$ parameters, which determine movement between the member classes. It follows that, in a scenario with equal amount of negative and positive media for each candidate, the most likely transitions will be a member switching and Democrat recruitment. Republican voters, on the other hand, are not likely to change parties or decay due to media influence (since $\sigma_1$ and $\delta_1$...
are below $10^{-5}$). When compared to the equivalent processes of decay and switching in the Democratic party, the probabilities could as well be neglected (smaller by about four orders of magnitude). Notice that these precise transitions in the Democrat party have the most impact on the outcome: $\sigma_2$ and $\delta_2$ have the larger sensitivity values, which means that the most effective way to change the result is by making Democrat voters more likely to respond to coverage against Obama and change class. Of course, this kind of change is difficult to actually achieve, but it can be addressed by changing the media content to target specifically Democrat voters and cause a bad image of Obama based on issues that Democrats care about.

Figure 10: Media diagram for values and sensitivity in the 2012 election: solid arrows represent transitions with higher parameter values ($> 10^{-1}$), dotted arrows correspond to lowest parameter values ($< 10^{-4}$) and line-dot dashing is used for intermediate values. Sensitivity of the outcome is visualized by color: from higher to lower sensitivity value, arrows go from black ($> 0.5$) to white tip ($< 10^{-2}$).
The general qualitative behavior of the system can be observed in Figure 10. A first conclusion is that a given amount of coverage causes more movement inside the Democratic party than inside the Republican one. It also appears that Democrat voters are more likely than Republicans to switch candidates because of negative media coverage.

3.2.2 2016 Election

The weekly poll data for the 2016 election (from the third week of May to the second week of July) are best fit by the lines in Figure 11 with the parameter values given in Table 4. Notice that media favoring Clinton is consistently greater than media favoring Trump, which is reflected in the increasing behavior of the Democrat voters’ fit. Correspondingly, the fit for Trump voters appears to be monotonically decreasing.

The parameters of higher value for this data set are the $\kappa$’s, which reveals a strong effect of media coverage in transitions between party members. The value of $\kappa_1$ being slightly greater indicates that, given the same amount of positive and negative coverage for both candidates, Republican members are more likely to become members of the other party than Democrat members.
Figure 11: Media model and best fit curves for the 2016 election: The plot on the top shows the modeled voting classes, $V_1$ and $V_2$; the middle plot presents the modeled member classes, $M_1$ and $M_2$; the bottom plot displays the amount of media coverage at a given time. For each candidate, we present the percentage of negative coverage affecting him and the added percentage of his own negative coverage plus the positive coverage for the competitor.
Table 4: **Parameter estimates and sensitivities for the media model applied to the 2016 election.** The SSE for the estimation is 0.0039.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Sensitivity</th>
<th>Parameter</th>
<th>Estimate</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_1$</td>
<td>3.66e-2</td>
<td>-0.2171</td>
<td>$\delta_2$</td>
<td>7.95e-4</td>
<td>0.0038</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>4.2e-3</td>
<td>0.0081</td>
<td>$\gamma_2$</td>
<td>8.02e-2</td>
<td>-0.2023</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>2.3e-3</td>
<td>-0.0053</td>
<td>$\lambda_2$</td>
<td>3.5e-3</td>
<td>0.0081</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>8.81e-4</td>
<td>-0.0102</td>
<td>$\sigma_2$</td>
<td>5.98e-4</td>
<td>0.0061</td>
</tr>
<tr>
<td>$\kappa_1$</td>
<td>8.48e-1</td>
<td>-0.0814</td>
<td>$\kappa_2$</td>
<td>8.06e-1</td>
<td>0.0717</td>
</tr>
</tbody>
</table>

Figure 12: **Media diagram for values and sensitivity in the 2016 election:** solid arrows represent transitions with higher parameter values ($> 10^{-1}$), dotted arrows correspond to lowest parameter values ($< 10^{-4}$) and line-dot dashing is used for intermediate values. Sensitivity of the outcome is visualized by color: from higher to lower sensitivity value, arrows go from black ($> 0.5$) to white tip ($< 10^{-2}$).

The transitions in which media have less effect are the direct movements between the voting classes. Keep in mind that the model assumes that this transition occurs solely
because of the amount of negative coverage of a candidate and it’s not dependent on the amount of positive coverage the other candidate is receiving (see Figure 3). In other words, negative coverage about the candidates is not very likely to drive voters to vote for the other candidate. In absence of positive coverage, for example, a given equal amount of negative coverage for both candidates is roughly a thousand times more likely to cause movement between member classes than between voting classes. Notice we do not mean *net movement*, which may be zero by cancellation.

Inside the parties, the media seems to have asymmetric effects. In the Republican party, for example, media coverage affects the decaying (from Republican voter to member) more than the recruitment (from Republican member to voter). Since media influence is the only factor affecting the transitions in the model, at any given time we observe a probability of decay (choosing not to vote) of approximately 10 times greater than the recruitment probability. In the Democrat party, however, the probability of recruitment is more than 10 times greater than the probability of decaying.
Figure 13: Media model and bar plot for the contrast of parameter sensitivity to the parameter value for the 2016 election: Y-axis is in log-scale.

4 Conclusions

4.1 Comparison of the 2012 and 2016 presidential elections using the person–to–person interaction model

In the context of person-to-person interaction, the level of influence by a group is qualitatively defined to be its constituents’ ability to cause a change-of-group in an individual upon contact (regardless of the direction of influence).

Using Figure 14 as a reference, we observe that the voter groups are the driving forces behind the 2016 election. On the other hand, the 2012 election is characterized by a share
Figure 14: **Person-to-Person interaction maps comparison**: 2012 (left) and 2016 (right) accounting for fitted parameters. Curved dashed lines represent interactions. Dotted lines, dashed lines and solid lines represent the transition with probability less than 0.3, 0.5 and 1, respectively. If a solid line represents the only possible result, then that transition has probability of 1.

in driving power between all groups with the voter groups having a slightly bigger effect on the system than the non-voting groups. In particular, during the 2012 election, the influence of the Democratic voters, $V_2$, is weaker in comparison to the Republican voters, $V_1$. This characteristic is coupled with the observation that the voter groups are more likely to be persuaded to switch groups by the non-voting members of the party they support by one another. Interestingly, this changes completely in the 2016 election, when
these voting groups are impartial to the influence of the non-voting groups, yet they are significantly influenced by one another. We propose the following interpretation of what these characteristics mean in term of the 2012 and 2016 elections.

For the 2012 presidential election, the interaction model shows that an interaction between a voting group and a non-voting group generally results in a loss of constituents for the voting group. This suggests that the voter turnout in the 2012 election should be less than the usual trends, regardless of the outcome of the election. This is consistent with what was observed in 2012. The voter turn-out rate with respect to the total eligible voters in 2012 was 57.5%, lower than in 2008 (62.3%) and 2004 (60.4%) [26].

In contrast, we observe different dynamics behind the current 2016 election. The election is driven by the influence of the voting groups. Specifically, the Republican voters have a significant influence on the Democrat voters, while they have no influence over the non-voting Democrat members. Democrat voters don’t have influence on other Democrat voters (they don’t repel them) and they do drive members of both parties to vote Democrat. This can be seen from the number of possible outcomes with significant probability (greater than 0.3) of happening after a Democratic voter meets a Republican voter, and vice versa. This number is two for the 2012 election and five for the 2016 election.

Note that the Democratic voters have a slightly weaker influence over the Republican candidate supporters compared to the influence that the Republican voters have over the Democratic voters. Yet, in 2012, Democratic voters exhibit absolute influence over Republican members to become Democratic voters, and, in 2016, they have this influence
on both member groups. This leads to the following speculation: members of the two
groups in general do not like the 2016 Republican presidential candidate, so in the event
that someone tries to convince them to vote against that candidate, they are likely to do
so. This is partially supported because it has been reported that regardless of Trump’s
current success in the presidential race, 67% of Americans view him at least somewhat
unfavorably (compared to 52% for Clinton) [5]. Note that this is purely within the context
of the person–to–person interaction model. On the other hand, the influence of the non-
voting groups in the 2016 election is weaker compared to their influence in the 2012
election. In fact, in the 2012 election $M_1$ and $M_2$ can influence $M_1$, $M_2$, and even $V_1$, and
$V_2$ to change their group affiliation. But in this election, their influence is limited to only
the non-voting groups.
4.2 Comparison of the 2012 and 2016 presidential elections using the media model

Figure 15: Media influence maps comparison: 2012 (left) and 2016 (right) accounting for fitted parameters. Solid arrows represent transitions with higher parameter values ($>10^{-1}$), dotted arrows correspond to lowest parameter values ($<10^{-3}$) and line-dot dashing is used for intermediate values.

As for the media effect on voting behaviors, there are some subtle but important changes between the two elections. First, note that overall, the media influence is noticeably more effective among the Republican members and voters in the 2016 presidential election than it was in the 2012 presidential election. This observation comes directly from comparing the magnitude of the parameters with index of 1 (movements from the Republican members and voters) between the two elections. Among the Democratic members and voters, the overall media influence is slightly weakened the same in the 2016 election compared to the 2012 election, e.g. three parameters decrease and two parameters increase. Yet, the increase in sensitivity toward the media among the Republicans is minimal. The only
parameter that increases significantly is the probability that voters stop supporting the Republican candidate ($\delta_1$). Note that this is not equivalent to stating that the 2016 Republican candidate is losing because we only consider members of the two parties and there is a significant percentage of voters, who do not belong to either party [51]. However, our speculation is that the huge amount of negativity that the media has on Trump is causing Trump supporters, or probably more precisely, longtime Republican voters, to stop supporting the Republican candidate.

Unlike the person-to-person interaction model where about half of the parameters have a significant value (larger than 0.05), only three of the parameters for the media model are significant ($\kappa_1, \kappa_2$ and $\gamma_2$). The significance of these values means that the nonvoting members of both parties are more susceptible to the media influence. The difference is that the Republican members do not generally decide to vote because of media influence. They simply switch sides to identify with the other party, whereas the Democratic members either exhibit a similar behavior or simply decide to vote for the Democratic candidate. The high value of $\kappa$ is likely because the media tends to not actively encourage people to vote. Instead, it simply spreads news. On the other hand, the significance in value of $\gamma_2$ is probably caused by something else, since it is not consistent in both elections. With respect to some of the smaller parameters, we notice that in 2012, there is a considerable tendency ($\sigma_2$ and $\delta_2$) for the Democrat voters to switch sides to vote for the Republican party or become non-voting members in response to negative news about the Democratic candidate. This observation is partially supported by the event when voters for Obama
plummeted after he was perceived to have lost the first debate to Romney in October 2012 [4, 27].

Furthermore, notice that the slight decrease in the sensitivity of the Democratic voters and members in 2016 is evident. This is because regardless of the huge amount of news covering the case of Clinton’s email-handling, over two thirds of Democrats do not seem to think of it as a major problem compared to a near unanimous disapproval among the Republicans, who think Clinton should have been charged for her mishandling of the confidential emails [21, 56]. Yet, the opposite effect happens for the Republican candidate, e.g. Trump is driving Republicans away due to the negative media coverage that he has been inducing[7]. Understanding this effect is critical since the media is always an important source of information for voters—especially since the 2016 election is marked with multiple controversial and complex issues such as the economy, immigration and terrorism [13, 23].

4.3 Conclusion

In conclusion, the media and interaction models reveal some general aspects of the 2012 election. In both cases, the parameter values defines which transition is driving the system. On the other hand, the sensitivity of the parameters helps define key strategies to alter the outcome of the election. Using sensitivity analysis and parameter fitting alone suggests that the most effective strategy for a candidate to win the election would be to focus on getting voters of the opposite party to vote for himself, instead of recruiting members from
either party. This makes sense theoretically, since a transition of voters directly affects one party positively and the other negatively and the member groups are small relative to the voter groups making the sensitivity analysis ineffective. One way that this strategy is carried out during an election is by attacking the candidate’s credibility/trustworthiness, for instance, Clinton’s careless handling of the emails as Secretary of State during the second term of President Obama or the candidate’s personal view, Trump’s stance on sensitive topics such as immigration.

Both models were able to fit the data well, where the fit using the media model gave a smaller error than the fit using the person-to-person model. The difference, however, is insignificant and could have easily been caused by artificial effects of parameter fitting and random errors. Thus, we refrain from concluding which best describes the most important factor in each election. Instead, we suggest a more thorough analysis to be done with more consistent data sets and similar numbers of parameters in both models.

The key elements of both models are the following:

- Media is most responsible for the transition between members of the two parties in both elections. The influence of media seems to be insignificant in all other transitions with one exception in 2012, which is the value of $\gamma_2$.

- The person–to–person interaction model suggests that the voter groups drive the 2016 election. This holds for the 2012 election but at a lesser extent.

- The person–to–person interaction model suggests that members in both party gen-
erally do not like Donald Trump. This comes from the observation that it is easier to influence them to change their minds to vote against Trump and the increase in the effectiveness of media influence among Republicans in 2016 (note that Trump has high amount of negative coverage).

Theoretically, the transition that is most sensitive to the election outcome is between voters of the two groups. Thus convincing voters to switch side should be the strategy to win the election. Yet, this is not feasible realistically. The key notes suggest that the most effective way realistically to drive the 2016 election is to manage the unfavorable opinions of members of both parties about Trump in their favor.

5 Limitation and future work

5.1 A note on the sensitivity values of parameters

As a side note, we would like to know what actions can effectively change the parameter with the highest sensitivity value. Unfortunately, this information is not directly obtainable with our models. This is due to each parameter being the net effective cause of the corresponding transition, so there is not a direct way to figure out the separate mechanisms that determine the value of the parameter. In order to suggest a method to obtain the mechanism behind the each parameter, here is an interesting observation. If we define an extreme parameter to be parameter with value either above 0.75 or below 0.25, then the person-to-person interaction model have 20 extreme parameters in 2012 and 15 extreme
parameters in 2016. We propose the following explanation for this observation.

Notice that our data for the 2012 election is first taken approximately three months before the election day and spans for 14 weeks, whereas our data for the 2016 election is first taken approximately six months before the election day and spans for 8 weeks. This difference in time period in which the two data sets are taken may indicate the different level of uncertainty in the decision people of the two parties make. Suppose this is not caused by the inherent differences between the two elections—which we do not know for certain. Then since there are significant difference between the two time period, e.g., the choice of vice president, the resulting difference in the number of extreme parameters can alternatively be explained by the difference in the stages of the election. If so, this suggests that a more appropriate way to model the dynamics behind voting behaviors is to represent the parameters using functions which take intermediate values initially, then as the election draws closer, these values become more extreme, e.g. greater than 0.75 or smaller than 0.25. This is making the assumption that people are less likely to change their mind as the election draw closer. A similar assumption has been examined before by Hau in his agent-based model on social effects in political elections [20]. Furthermore, since we are interested in knowing what mechanism drives change in parameters, we can structure these functions to incorporate different assumptions and use data fitting as a way to justify which assumption is appropriate. This should be carried out after determining which parameters matter most in the election under consideration.
5.2 Limitations and Future Work

There are two major sets of limitations in our work. The first corresponds to the assumptions of our models and the second is related to our method and the available data.

The political voting system is extremely complex, and naturally the presented models are only rough approximations of reality in order to measure the fundamental driving forces behind a presidential election. Applied to the 2012 and 2016 elections, sometimes, our models gave insights into the underlying mechanisms behind each election, which is supported by data. Other times, it gave suggestions that seem inconsistent with reality.

To construct our models, we make many strong assumptions such as the assumptions of homogeneous mixing of the population and the assumptions that voting behaviors can only be changed through interaction with another person or by the influence of the media. These strong assumptions enable us to focus our analysis on the effect of social media and person-to-person interaction on voting behaviors in a presidential election.

The second limitation is our collection of data. Aside from the random error that is inherent to polling data, we also use data sets from multiple sources for fitting of the 2012 election. This can potentially have major impact on our analysis. Yet, in our case, the effect seems to be minimal judging from the fitting errors and the trends of the data points. Another limitation associated with the data comes from our use of the data to find the parameters for the models. These polling data take a sizable amount of time to collect. But when fitting, we assume them to be taken exactly at one point and two
consecutive polls are always exactly seven days apart. This assumption is necessary for our discrete-time Markov model, but it is another potential significant source of error. Not only quantitatively but philosophically, it is a strong assumption to consider that the least amount of time in which a person can change their political position is a week. Furthermore, to ease the complexity of the model, we choose to approximate the media coverage to be constant between any two data points. In addition to that, since our models, especially the person-to-person interaction model, have a large set of parameters, it is easy to over fit. This can be addressed using a more thorough sensitivity analysis to find the least sensitive parameters for each election and eliminate them from the model. This can be carried out using the Latin Hypercube Sampling (LHS) method to vary all parameters within a certain range simultaneously to find their effects on the system relative to one another.

As mentioned previously, our use of the data mainly bases on the number of self-identified likely voters for each party. This is appropriate for our purpose; however, for a more thorough analysis of the underlying social and media effects, a more detailed data set could be used. For example, a data set with information on the ideology of the likely voters can help us better identify theoretically the different directions of influences (the arrows in the graphs). We also notice that the error for fitting the data collected from a single source is significantly smaller than the one fitting data collected using multiple sources. This is understandable due to the inconsistency between differences in methods for data collection and data samples. A better collection of data would be to use the
average from multiple weekly sets of data, since it is a trend that the average of pollsters tends to be a good predictor for the outcome of the election. This, however, is difficult to achieve because polling data is generally collected through private agencies. Moreover, as mentioned in the previous section, data sets taken at different time period with respect to the election day could potentially cause ambiguity in the meaning of the parameters. Thus, for the purpose of comparison of voting behaviors between different election, data sets from the same time period (relative to the election day) should be used. On the other hand, if the purpose is to find the mechanisms that determine the value of the parameters, then a data set of significant length, e.g., a data set that spans six months before the election day, should be used in parallel with a function forms for the parameters.

Additionally, there is a significant proportion of eligible voters who are not aligned with either major party, thus, they are excluded from the data used to fit our models. This is important in two aspects. First, our models cannot measure the general trend of the voting behavior. Secondly, without the consideration of this population, the two groups that do not wish to vote for either candidate, $M_1$ and $M_2$, are much smaller than in reality. This may limit the effectiveness of the sensitivity analysis and render any conclusion drawn from it meaningless.

Recall that our fitting does not guarantee a global minimum for the fitting error. However, since we do not have a complete understanding of the distribution of the parameter values with respect to the fitting error, or the distribution of local minima, we cannot say for sure what effects it could have on our system. Furthermore, obtaining the global
minimum for the fitting error does not mean we have the most realistic set of parameters. This can easily be some artificial properties of the model or our numerical methods. Thus a set of parameters with global fitting error may not help us explaining reality better. We do not disregard the potential benefits of having a global minimum fitting error, we simply acknowledge the possibility that it can be an artifact of our method. In fact, since we vary our initial guesses randomly each time we run the program for parameter fitting and still obtain the same fitting error, we may as well assume that our fitting error is the global minimum on a large region in the parameter space.

Lastly, other than using more accurate data and methods to analyze individual elections, further work should be done on studying how elections affect voting dynamics in the future. Our framework could be expanded towards comparing voting behavior across multiple campaign elections, incorporating a cascade effect [16] in which campaigning efforts in a given year can be carried on to the next election and influence the voting population. One last note on our framework: is it really appropriate to model voting behaviors using a Markov chain? Both models assume what happens in the future is only dependent of what is going on in the present. This is arguable because a decision can be viewed as the output of a collection of information. Thus, we also hope to extend this framework to a non-memoryless process.
6 Acknowledgement

We would like to thank the Mathematical and Theoretical Biology Institute (MTBI) co-Directors Dr. Carlos Castillo-Chavez, Dr. Anuj Mubayi, and Dr. Marlio Paredes for giving us the opportunity to participate in this research program. We would also like to thank Associate Director Sherry Woodley and Coordinator Ciera Duran for their efforts in planning and executing the day to day activities of MTBI. We also want to give special thanks to Daniel Burkow and Dr. Leon Arriola for his guidance and constant support.

The research has been carried out at MTBI which is a Research Experience for Undergraduate (REU) summer program at the Simon A. Levin Mathematical, Computational and Modeling Sciences Center (SAL MCMSC) at Arizona State University (ASU). This project has been partially supported by grants from the National Science Foundation (DMS1263374), the Office of the President of ASU, and the Office of the Provost at ASU.

References


