Using Agent-Based Modeling to Understand the Impact of Community Interactions on Voter Apathy and Election Outcome

G.A. Brophy, A.R. Rips-Goodwin,
L.A. Wilson, and A.L. Lewis

Abstract - Voter turnout in U.S. elections tends to fluctuate, with multiple factors influencing a prospective voter's participation. In this investigation, we focus on how a registered voter might change their inclination to participate in an election based on their perception of others' political views within their community. We present an agent-based model (ABM) designed to explore community-based driving mechanisms of voter turnout and election outcomes. By conducting a sensitivity analysis, we examine which of our model inputs—including factors such as party-affiliation breakdown or the degree to which voters tend to communicate in echo chambers—have the greatest impact on voter turnout and the extent to which election results accurately represent the majority opinion in the electorate. We find that within our model framework, turnout rates depend heavily on party affiliation heterogeneity as well as how strongly a voter is influenced to vote or abstain by their politically-based interactions within their community. Additionally, we find that it is possible for an unexpected winner to emerge victorious due to an increase in mobilization among minority party members under certain conditions.

Keywords : agent-based modeling; sensitivity analysis; voter apathy

Mathematics Subject Classification (2020) **:** 91B12

1 Introduction

Controversial candidates and an increasingly polarized political system have fueled a recent surge in voter turnout across elections in the United States. 62.8% of the voting age population participated in the 2020 presidential election, with similar surges in turnout observed in modern midterm elections. However, the U.S. still trails many other countries in the Organization for Economic Cooperation and Development (an organization of countries committed to democratic values working to develop and coordinate economic and social policies) in terms of voter engagement, ranking 31st among 50 countries for voter turnout among voting age citizens in recent national elections [1]. As such, much research has been conducted in an effort to understand what factors drive voters to the polls. In this investigation, we seek to answer questions about community influences on voter apathy through the use of an agent-based modeling framework.

Agent-based models (ABMs) are simulations in which systems are modeled as a collection of autonomous individuals, or "agents" [2]. ABMs allow for a number of design concepts that are typically difficult to incorporate under more homogeneity-based frameworks, such as systems of differential equations. For example, ABMs treat individual agents as autonomous entities, who make independent decisions based on uniquely assigned parameter values. Agents may possess the ability to interact with other agents, sense each other's characteristics, and/or adapt their behavior in response to what they sense. ABMs can easily incorporate stochastic elements, allowing for random assignment of parameter values from a given distribution, or random choice between specified behavioral options. In general, individual-based simulations that allow for rich heterogeneity—like ABMs—often permit more realistic representations of complex phenomena as compared to other types of traditional mathematical modeling [3]. In particular, the complexity that defines and emerges from ABMs makes them well suited to model behaviorally charged problems such as voting and voter apathy. Our use of an ABM framework allows us to target how an individual voter's interactions within their community might lead them to make a particular choice with regard to their participation in an election, namely by controlling for certain factors and allowing others to vary at the individual level.

Within the vast world of election studies, agent-based models provide a means to juxtapose historical data with hypothesized phenomena described in the literature. A rich amount of political research focuses on voter motivation, voter apathy, and voter turnout through sociological, psychological, and technological lenses, while the genre's agent-based models tend to examine the implications of a specific behavior or sociological problem [4, 5, 6, 7]. Instances of this include using agent-based models to examine the impact of candidate campaign efforts and individual voter behaviors, alongside efforts to better understand external drivers of voting [8, 9, 10, 11]. Other research uses the advantages of agent-based modeling to gain a better understanding of different voting systems [12, 13], or to examine voter turnout through the lens of voter preference and party dynamics [14, 8, 15, 16, 17]. Our work contrasts these existing simulations by fixing assumptions about what motivates voters, and altering the extent to which these drivers are more community- versus internally-oriented as a means of studying voter turnout and election results, rather than exploring how electorates behave in the presence of new factors and political theories. Additionally, compared to several other ABMs that model voting, our results are not bound to the conditions of a singular election or voting population, allowing for possible adaptations of our work to such specifications in the future.

The social hypotheses underscoring several political agent-based models explore both the types of social interaction and behavior tied to voting patterns and the role of new factors like increased social media presence [18, 19, 7]. Alongside the substantial body of work detailing the connections between personal identity and propensity to vote, assumptions about voter motivation are often fueled by research detailing the underlying patterns tied to voter turnout [20, 21, 22, 23]. Within this sub-field, the influence of community is recognized as a contributing factor [24, 25]. Our model highlights this influence in comparison to voters' unchangeable demographic characteristics in order to pinpoint the impacts of political interaction in a way currently underrepresented in the literature. With the benefits of agent-based modeling and our exploratory lens, we contextualize previous work with a broader perspective on the connections between voter turnout, election results, and a voter's perception of the opinions of their fellow voters.

Several theories have been applied to investigate how community effects impact voter turnout. One such framework is rational choice theory, which hypothesizes that prospective voters will turn out at the polls if they think their vote could be pivotal, rather than being driven by the actual practice of voting. However, this theory succumbs to the "paradox of voter turnout," which describes the underestimation of turnout in large elections [26]. We turn to an alternate explanation posed for the variability in voter turnout: a laboratory study conducted by Levine and Palfrey (2007) that explores the comparative static effects, or predictors of the impacts of external factors, of the rational choice model [26]. Two phenomena observed in this study are particularly relevant. The first is the "Competition Effect," which represents the idea that voter turnout should be higher in elections that would be highly contested if all eligible votes were cast. The second is the "Underdog Effect," which details that turnout for a less popular candidate should exceed that for a more popular candidate. The study found strong evidence that each of these two strategic responses arise under varying experimental conditions [26]. As such, we model voter behavior under the influence of these hypotheses separately to consider how each may impact turnout. In effect, our model imposes the behaviors observed as results in Levine and Palfrey's study, reversing the direction of analysis by looking at how community-driven factors may influence voter behavior under each of these trends.

While the ABM framework is well-suited to realistically model the nuances of voting behavior, it is less straightforward when it comes to interpreting results. In particular, calibrating the parameters and settings of agent-based models to existing data can be very difficult, since individual-level data is not often available, and mechanisms incorporated into the ABM may not map directly to observable outputs. For this investigation, we aim to keep our model generalizable, rather than estimating the parameters to align with a certain region or community. As such, this model is not intended to be used for predictive purposes at this time. Rather, this ABM serves as an opportunity to explore how various input settings can create a range of model outcomes, an essential first step towards creating an explanatory model of community-driven voting behavior. Since fixed values of the parameters are not prescribed, this approach necessitates a thorough assessment of the sensitivity of the model quantities of interest to each of the inputs. Moreover, because agent-based models have highly flexible functionality, conducting such analyses is often computationally expensive. Still, performing a robust sensitivity analysis is an imperative part of the model development process, as it enables us to determine which specific factors—or factors in combination—create conditions resulting in high levels of voter apathy. In this investigation, we follow the approach to sensitivity analysis detailed in Borgonovo, et. al [27]. This framework highlights four key components in an ABM-centric sensitivity analysis: factor prioritization, direction of change, interaction quantification, and robustness analysis. Addressing each of these four priorities leaves us better equipped to interpret our results and answer our driving questions regarding voting behavior.

In Section 2, we outline our model development process using the "Overview, Design concepts, and Details" (ODD) framework for agent-based models [3, 28], and then summarize the metrics that we will use to perform our sensitivity analysis. Results are presented in Section 3, where we study the impacts of each of our community-based inputs on two quantities of interest: overall voter turnout, as well as the discrepancy in proportion of votes that are attributed to each candidate as compared to the party affiliations of the electorate. Finally, we summarize our findings and discuss potential future directions for research in Section 4.

2 Methods

2.1 Overview, Design Concepts, and Details (ODD) Protocol

In this section, we describe our model construction and methodology using the "Overview, Design concepts, and Details" (ODD) protocol for agent-based models proposed by Railsback and Grimm [3, 28]. The ODD protocol is composed of seven constructs: Purpose and patterns; Entities, state variables, and scales; Process overview and scheduling; Design concepts; Initialization; Input data; and Submodels. These seven elements serve to concisely describe the ABM independent of the algorithms and programs used to execute it. The implementation of a standardized model description like the ODD promotes communication among scientists from different backgrounds, giving the ABM greater utility and reproducibility [28]. We employ this framework in order to conform to standard practices for describing components of agent-based models—which often require elements absent in other types of models—at a level enabling replication by other researchers. Figure 1 gives a summary of the ODD protocol, with components that are relevant to our particular model in bold font. Each of these components are described in detail in the following sections.

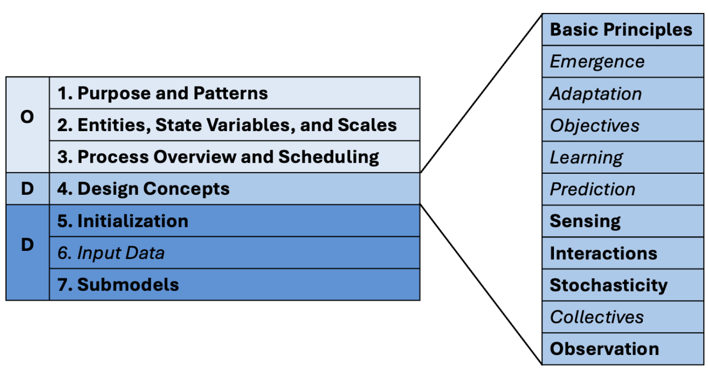


Figure 1: Summary of the Overview, Design concepts, and Details (ODD) protocol. Components that are included in our model are listed in bold font. Figure is adapted from [28].

2.1.1 Purpose and Patterns

With this model, we aim to understand how an eligible voter's perception of the political climate in their community may impact their decision to participate in an election. Specific questions that we wish to address include:

- Under what types of community conditions is voter turnout maximized?
- Under what types of community conditions do we observe an election outcome that is not representative of the majority opinion in the electorate?
- How do different levels of activism displayed by voters who perceive themselves to be in the minority party (based on their surrounding neighbors), impact both voter turnout and election outcome?

By studying a variety of different conditions which affect community political make-up (e.g., the amount of diversity in political opinion among neighbors, the degree to which agents "cluster" with like-minded agents, the size of agents' social circles, etc.), we illustrate how these conditions may coalesce to generate various different election outcomes.

Our model represents a bipartisan political system where voters are classified as members of either the "blue" or "red" parties. While the model construction is loosely based upon the two-party system of the United States—with Democratic and Republican parties, respectively—the model in its current form could be easily adapted to represent any other two-party political structure. Importantly, the model setup is not intended at this time to represent one specific area or population within the US, but rather to symbolize a random selection of eligible U.S. voters.

From past elections in the United States, we know that voter turnout can vary drastically among different communities, even in national elections that seem to affect all communities such as presidential elections. Likewise, with each major election, we see districts in which an underdog candidate takes the vote, despite the majority of the electorate expressing a preference for a different candidate. Our model is intentionally designed to generate results that are consistent with a wide range of outcomes observed in national elections at the state and district levels—though calibration to specific data is not yet possible, selecting our parameter ranges and settings in a way that allows us to reproduce known results is an essential first step. Furthermore, such flexibility in the model construction process allows us to examine which community factors may be driving such discrepancies in election outcome.

2.1.2 Entities, State Variables, and Scales

The following entities are present in our model: agents (representing eligible voters), grid cells (representing agent "locations" within their community), and the environment, which controls system-level simulation settings. An agent's "location" on the grid does not correspond to a geographical position. Rather, it is a randomly assigned position based on political affiliation that will be used to define an agent's "social circle" within the larger community, that is, within the set of all agents appearing on the grid. In all, 2,601 agents (situated upon 2,601 grid cells), are arrayed in a 51×51 grid, with each agent residing at the center of their assigned grid cell. While the relative size of our agent population is similar to that of various smaller communities like counties or neighborhoods within the United States, this ABM is not intended to simulate a specific region or group. Thus, its population size is chosen arbitrarily, with its grid size accommodating our chosen number of voters. Agents remain stationary throughout the simulation, maintaining the equally-spaced separation of agents within the grid view. We elaborate on our simulation of community interaction in Section 2.1.4.

Each agent possesses several starting attributes that distinguish it from other agents. As the simulation progresses, other state variables—to be described later—are calculated and assigned based on these starting quantities.

The first of the agent state variables, `party`, represents the strength of the agent's political affiliation, which is important both to establish an agent's baseline probability of voting without considering community factors as well as their potential influence upon surrounding agents. Values are assigned on a continuous scale from 0 to 1, with 0 representing the strongest possible "blue" affiliation, 1 representing the strongest possible "red" affiliation, and 0.5 indicating a truly moderate voter. Agents are further categorized by their party assignments, indicated by `party-cat`, such that all agents with `party` values in the 0-0.5 range are assigned `party-cat =` "blue," and all other agents are assigned `party-cat =` "red." In order to illustrate the diversity of political opinions in the user interface, each agent's grid-cell is assigned a `pcolor` value, which is used to color the cell according to the category and strength of the agent's political affiliation. In the model interface, a dark blue grid-cell corresponds to a `party` value close to 0, white represents a truly moderate voter with a `party` value near 0.5, and dark red indicates a `party` value close to 1.

Each agent is also assigned a `radius`—with possible radii values being dictated by an environmental parameter `mean-radius`—which defines the size of their individual social circle. That is, the magnitude of an agent's `radius` dictates the number of neighboring agents with whom they will interact throughout the simulation. Allowing for agents to have social circles of varying sizes enables us to consider whether the size of one's social group may affect the degree to which community-based factors influence their decision to vote.

Finally, each agent is assigned a `demo-prob`—a probability that they will vote in the election based upon their demographic characteristics or individual identity—as a baseline value from which their eventual voting probability will be derived. The inclusion of this `demo-prob` variable is intended to help control for individual-based factors that may play a vital role in determining an agent's likelihood of voting—such as their education level, party affiliation, socioeconomic status, age, race, sex, etc.—allowing us to focus specifically on the community-based factors that we wish to investigate. These `demo-prob` values are generated using an external logistic regression model calibrated to data from a 2016 U.S. election survey [29], to be described in Section 2.1.5. The fixing of these demographically-based voting probabilities in advance of the simulation is conducted in order to control for individualized characteristics that might alter one's voting behavior, as our focus throughout this investigation is solely based on how *community*-based factors contribute to voter apathy. Thus, in each simulation we consider snapshots of an electorate, where agent demographics and social networks remain fixed. We enforce such a framework to both reflect the limited timescale of our model and to account for the impact of individual characteristics while focusing primarily on community interaction in this study. By rendering personal demographics static yet dependent on real voter data, we account for the complex relationships between personal lived experience and likelihood to vote in a computationally simplistic manner. Links between personal characteristics and different voting behavior are represented by differing demographics-based voting probabilities. We assume that our electorate represents a random sample of the U.S. voter population over a temporal period of months leading up to an election. Hence, we reflect both the stability of our demographic factors of choice (i.e. age, race, and education level) within such a time period while allowing them to exert influence on individual decisions in accordance with real data.

The state variables assigned at the start of simulation are summarized in Table 1. It should be noted that all of the settings controlled by the environment entity are static with respect to time, and are thus characterized as parameters rather than state variables. Further details about all state variable and parameter assignments will be provided in later sections of the ODD protocol.

The spatial component of our model is not defined in the traditional sense. An agent's `radius` variable dictates the set of neighbors with whom they will interact, but these interactions could represent both in-person as well as online social interactions. The `radius` values vary among agents so as to represent the differing sizes of social circles among voters. Both the size of the grid cells and the overall environment are left unspecified, as the view is not spatially bound to a real landscape.

Likewise, the model is not temporally bound to a specified length of time. The entire simulation takes place in one "time" step, during which each of the voters adjusts their likelihood of participating in the election based on interactions they have within their specified social circle. At the conclusion of this time step, the election is conducted. In this sense, the single time step is representative of a period of unspecified length right before an election, when voters are likely to be discussing politics with those in their social circles and forming a decision about whether or not they will participate in the upcoming election.

2.1.3 Process Overview and Scheduling

In this section, we provide an overview of the execution order of the ABM. Specific details about each of the submodels are provided in later sections—here, we simply discuss how each of the submodels are organized with respect to scheduling. The main scheduling overview is as follows:

1. **Setup phase:** Agent and grid cell starting state variables—`party`, `party-cat`, `radius`, `demo-prob`, and `pcolor`—are assigned. Environment parameters `comm-effect`, `a`, `party-split`, `mean-radius`, and `degree-clustered` are specified.

State Variables			
Name	**Entity**	**Description**	**Possible Values**
`party`	Agent	Strength of political affiliation	Continuous on $[0, 1]$
`party-cat`	Agent	Category of party assignment	`blue` or `red`
`radius`	Agent	Defines size of social circle	Continuous, with value dependent on `mean-radius` parameter
`demo-prob`	Agent	Inclination to vote based on individual characteristics	Continuous on $[0, 1]$
`pcolor`	Grid Cell	Visual indication of political affiliation	Dark blue, medium blue, light blue, white, light red, medium red, dark red.
Environmental Parameters			
Name	**Use in Model**	**Description**	**Possible Values**
`mean-radius`	Initialization	Defines average size of an agent's social circle	4, 5, or 6
`degree-clustered`	Initialization	Prob. agent locates near like-minded agent during setup	0.5, 0.51, 0.52, . . . , 0.99
`party-split`	Initialization	Categorical party breakdown of the electorate	"Strong Blue Majority," "Partial Blue Majority," "No Majority," "Partial Red Majority," or "Strong Red Majority"
`a`	Submodel	Factor weighting individual versus community-based likelihood to vote	0.1, 0.2, 0.3, . . . , 1.0
`comm-prob`	Submodel	Underlying hypothesis for voter behavior	"Competition Effect" or "Underdog Effect"

Table 1: Descriptions of state variables and environmental parameters.

We elaborate further on our environmental parameters in Section 2.1.4; a detailed summary can be found in Table 1. The user interface is generated according to the environment parameter choices.

2. **Agent-update phase:**

   (a) Agents execute the `check-neighbors` submodel and determine their `similarity` index, quantifying how similar the political beliefs of other agents within their social circle are to their own.

(b) Agents compute their `comm-prob` value—their probability of participating in the election based on community factors—by using their `similarity` index in combination with the `comm-effect` environment parameter.

(c) Agents create their `vote-prob` value, or overall voting probability, by combining their `demo-prob` and `comm-prob` values through a linear combination weighted by the environment parameter `a`.

3. **Election phase:** The election is conducted. Each agent "decides" whether or not to vote by comparing their assigned `vote-prob` value to a random number on the interval $[0, 1]$. For agents who decide to vote, their vote is cast in favor of the party assigned to them in the `party-cat` variable.

4. **Observation phase:** Results are tallied and displayed in the user interface, including overall voter turnout and the number of votes cast for each of the `blue` and `red` parties.

In all agent procedures, the agents execute the given command in a randomized order, since the order in which the agent variables are updated has no bearing on the variable assignments.

2.1.4 Design Concepts

In what follows, we discuss the major design concepts that appear in this ABM, in order to provide justification for all choices made during our design process. We discuss the five concepts which are relevant to our model, selected from the larger list of 11 common ABM concepts provided in [3, 28].

Basic Principles:

Our study of voter engagement versus voter apathy is based largely upon work conducted by Levine and Palfrey [26], in which a laboratory experiment was conducted to determine comparative effects in voter turnout. We highlight two of these effects for inclusion in our model: the "Competition Effect" and the "Underdog Effect."

The Competition Effect hypothesizes that prospective voters are most driven to vote when they perceive that the outcome of the upcoming election is likely to be close. We adapt this idea to our agent-based model by having voters define their perception of the "closeness" of the election based on the interactions that they have within their own social circle. For instance, a voter who believes themselves to be surrounded by voters of the same party as themselves will become apathetic, sensing that their vote will have no bearing on the outcome of the election. Likewise, a similar degree of apathy is observed among voters who perceive themselves to be surrounded almost entirely by members of the opposing party. However, if a voter perceives a nearly even split in party preference based upon their community interactions, they will be highly motivated to vote since the inclusion of their vote may well impact the final result.

In contrast, the Underdog Effect models an "activist" minority party, in which voters who believe they are in the minority among their community are highly motivated to vote in an effort to tip the election outcome toward the candidate whom they perceive to be the underdog. Similarly to the Competition Effect, voters who are surrounded by members of their own party (thus perceiving themselves to be in the majority) become apathetic in their voting behavior. In extreme cases, this imbalance—in which voter turnout is higher among members of the minority party than in the majority party—can have the effect of "flipping" the election outcome from the anticipated result.

The underlying hypothesis dictating voter behavior in our model is controlled by the `comm-effect` environment parameter. When running the model from our NetLogo interface, the user may select either `comm-effect` = "Competition Effect" or "Underdog Effect" during the setup phase. This choice then dictates how agents employ their `similarity` index to influence their probability of voting, to be described in Section 2.1.7.

Sensing:

Within our model, agents are able to sense the political affiliations of the other agents who reside within their predefined `radius`, as it is assumed that agents are likely to be engaging in political discussions with these neighboring agents. Moreover, agents are more strongly influenced by neighboring voters who have "extreme" opinions; i.e., values of the `party` variable close to 0 (strong blue) or 1 (strong red). When an agent computes their `similarity` index, which quantifies the degree to which agents within their circle express the same opinions as themselves, neighboring agents with more extreme opinions are weighted more heavily than those with more moderate affiliations. This choice was made in order to reflect the fact that voters with stronger party affiliations are likely to be more vocal in expressing their political views [30], and thus should carry more weight than moderate neighboring agents, who are more likely to keep their views private, with political affiliations that are therefore more difficult to sense. The specific details of this sensing mechanism are further described in Section 2.1.7.

Interactions:

A major strength of agent-based models is their ability to characterize agent interactions on a local level. In this ABM, agent interactions are defined at the local level using each agent's assigned `radius` value. The magnitude of this value translates into the size of the agent's "political social circle" within the larger community grid, which we we define to contain all of the individuals with which an agent discusses politics and political opinions. Using the `check-neighbors` submodel described in Section 2.1.7, agents interact with their political social circles while remaining in their assigned grid square. Social circle radii vary among agents, representing the fact that the number of politically-based interactions should vary drastically among different voters; some voters are very politically active within their community or online, while others prefer not to discuss politics with more than a handful of acquaintances. When making the decision about whether or not to participate in the election, each agent is influenced only by those other voters who reside inside their radius: i.e., those voters with whom they have interacted. Further details on this procedure are supplied in Section 2.1.7.

Stochasticity:

Another major benefit to using agent-based models is their ability to easily incorporate stochastic elements. In this model, we employ stochasticity to produce variation in our results while avoiding the complexity of modeling what drives such a difference. Specifically, we do this in several ways:

- The initial agent state variables `radius` and `party` are randomly assigned, within the constraints imposed by the user-selected parameters `party-split` and `mean-radius`. At this step in the model, the initialization of these attributes reflects their independence and ensures variation in initial conditions without influencing overall model trends.

- During the election phase, an agent's decision to vote is based upon the comparison of their `vote-prob` variable to a random number uniformly drawn from the interval $[0, 1]$. This approach aligns with a frequentist perspective of probability as a "long-term frequency"—for example, a voter with a probability of 0.73 would vote in approximately 73% of elections. Stochasticity in the election allows us to incorporate this view while maintaining simplicity in the submodel dictating the vote-counting procedure.

Observation:

As discussed in Section 2.1.1, our main objective for this model is to study voter turnout and election outcome under various different conditions. As such, for each simulation, we track the number of votes cast among our 2,601 agents, in addition to tracking how many votes were cast for each of the two parties. From this, an election outcome can be determined.

At the completion of each simulation, the winning party and their number of votes are displayed on the user interface. Additionally, we include a statement of the voter turnout, displayed as a percentage of total agents.

2.1.5 Initialization

We begin the initialization process by selecting the environment's `party-split` parameter, which dictates the balance between blue- and red-party voters in the simulation. The model is constructed to allow for five different options of this input: "Strong Blue Majority," "Partial Blue Majority," "No Majority," "Partial Red Majority," and "Strong Red Majority." To generate agent `party` distributions that align with each of these scenarios, we draw 2,601 `party` values at random from a beta distribution reflecting the desired `party-split`. Figure 2 illustrates how our sampled `party` values adhere to these distributions.

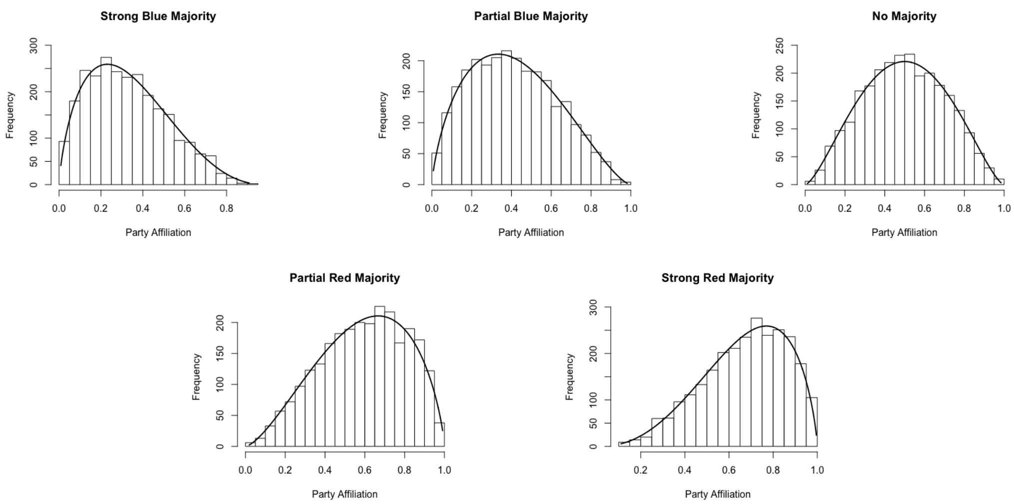


Figure 2: Party affiliation distributions for each of the five possible `party-split` scenarios.

The assignment of the agent state variable `demo-prob` is based upon input from an external model which was trained on data from the 2016 Cooperative Congressional Election Survey [29]. The survey recorded information about demographic characteristics, political ideology, and election participation for 64,600 respondents over a period of four months surrounding the 2016 U.S. general election. Because of evidence suggesting that inclination to vote is tied to the strength of one's political ideology [30]—among other factors—we train a logistic regression model on this data set with the goal of characterizing the relationship between an agents' `party` and `demo-prob` state variables and controlling for other demographic factors that might impact an individual's likelihood of participating in an election. Recall that an agent's `demo-prob`, or demographic-based voting probability, is intended to represent their baseline probability of participating in an election in the absence of community-based factors.

We employ the Akaike Information Criterion (AIC) to determine the ideal set of explanatory variables for predicting voter participation in the election across all possible sets of variables included in the 2016 CCES data set, with a goal of obtaining a model with strong predictive capability without overfitting the data. The optimized set of explanatory variables includes age, sex, race, income, level of education, home ownership status, immigration status, and party affiliation (ranked on an integer scale from 1-7, with 1 representing "Strong Democrat," 4 representing "Moderate," and 7 representing "Strong Republican"). The dependent variable is a binary indicator of whether or not the participant voted in the election. After responses indicating non-citizen status ($n = 1526$) and those with missing data ($n = 146$) are removed, a sample of 62,928 eligible voters remains, all of which are used to train the logistic regression model to yield a voting-probability output. These voting probabilities, or `demo-prob` values, range from 0.019-0.99, with a mean of 0.71, a median of 0.78, and a standard deviation of 0.22. We note that these sample statistics are likely skewed higher than those for the full U.S. population, since respondents electing to participate in the full survey would have been more likely to be politically engaged. The predicted likelihood of election participation is attached to each of the respondents in our data set, and saved for later use when sampling respondents to represent agents in our interface.

We then assign agents `party` values based on the specified `party-split`. Specifically, `party` values are sampled from a beta distribution parameterized by `party-split`, which produces continuous values on $[0, 1]$. These values are mapped to the discrete set $\{1, 2, \ldots, 7\}$, which corresponds to the survey's ideology rankings. We then select a survey respondent from the dataset with the same discrete party value and assign the respondent's predicted voting probability, or `demo-prob`, to the agent. This ensures that each agent's `demo-prob` aligns with the voting likelihood of a survey respondent who shares the same political views. The resulting `party` and `demo-prob` pairs for the 2,601 agents in each of our five `party-split` scenarios are saved into text files that are read into the model program.

After choosing the `party-split` setting—and the corresponding set of `party` and `demo-prob` pairs to be employed—the values of these variables are assigned to the agents in the grid view according to the environment's `degree-clustered` setting, which controls the extent to which agents tend to interact with like-minded individuals. The assignment is performed in the model's `create-clusters` submodel, which contains the following sub-routines:

Procedure for `create-clusters` Submodel:

1. `load-file-data`: Values for agents' `party` and `demo-prob` state variables are read in from the file corresponding to the user-defined choice of the `party-split` parameter.

2. `create-cluster-seeds`: 26 agents (1% of the total agent population) are selected at random as seeds for the clusters. These seed agents are assigned a `party` value (and corresponding `demo-prob` value) at random from the full list.

3. `initialize-party-subsets`: The remaining `party` values (with attached `demo-prob` values) are separated into two subsets based on their party affiliation: blue (for `party` values in the interval $[0, 0.5]$), and red (for all other `party` values).

4. `grow-clusters`: Agents who do not currently have a `party` assignment are prompted in a random order to consider their closest neighbor with a party affiliation. These agents then select their own party affiliation depending on the value of the `degree-clustered` parameter; specifically, the value of `degree-clustered` determines the probability with which the agent will select from the list that aligns their party affiliation with that of their closest affiliated neighbor. That is, if `degree-clustered` $= 0.5$, agents will choose their party affiliation entirely at random; as `degree-clustered` increases towards 1, the agent is more and more likely to choose a party affiliation that matches that of their closest affiliated neighbor. Within the selected list, the agent's `party` value is selected at random, if any are remaining.

5. assign-remainder: When either of the red or blue subset lists are exhausted, the assign-remainder sub-routine is triggered, assigning party values (and corresponding demo-prob values) to any remaining unaffiliated agents at random.

6. assign-color: Finally, the grid cell for each agent is colored according to its agent's party value. Though the party values are continuous, grid colors are assigned according to a discrete set of seven color options, referred to in the code as pcolor.

Upon completion of the create-clusters protocol, each agent is assigned the party-cat designation corresponding to their party value: blue if party is in $[0, 0.5]$, and red otherwise. These designations will be used in the final election to determine vote counts for each party.

The final step of the initialization process is to assign radii to each of the agents, which will dictate how many different interactions they each have. The user-defined environment parameter mean-radius controls the average interaction radius for the agents. This value may be set to discrete choices of 4, 5, or 6 (corresponding to roughly 52, 80, or 114 other agents contained within a given agent's social circle, respectively). The size of these social circles reflects the modern impact of social media and online discourse, increasing the everyday voter's exposure to political ideas and sentiments within their community. Furthermore, our simulated social circles are analogous to larger communities discussed in Rolfe's *Voter Turnout: A Social Theory of Political Participation*. Interaction circles with the lowest setting mimic the larger of potential "regular contact" networks, while social circles from the latter two options for mean-radius surmise groups of people whose connections do not mandate a large degree of emotional closeness [31]. This suits our community framework well, as people who interact politically do not necessarily need to have a previously-established social rapport in-person or online, and allows us to draw conclusions based off of larger, looser interactions (as opposed to analysis of person-to-person impact). To allow for various agents to have social circles of differing size, we vary the size of the interaction radii by assigning agents' radius state variables randomly from a normal distribution centered at mean-radius with standard deviation of $\frac{2}{3}$ during the setup phase. This choice of standard deviation ensures that nearly all of the agents (that is, roughly 99.7%, according to the empirical rule) will have radii within the values mean-radius± 2, a range chosen to ensure that all agents have a positive number of interactions even with our smallest mean-radius setting.

2.1.6 Input Data

The only external data utilized in this simulation is used in the initialization stage to assign demo-prob values to agents; this procedure was discussed in Section 2.1.5.

2.1.7 Submodels

The model code contains three major submodels: create-clusters, check-neighbors, and elect. The create-clusters submodel was described in the initialization procedures in Section 2.1.5. Here, we provide additional details regarding the other two procedures.

Procedure for `check-neighbors` Submodel:

1. Agents are selected in a random order to gather information about their neighbors. For a given agent, let $i = 1, 2, \ldots, n$ represent the numbers of the n other agents in their social circle, as defined by their `radius` value.

   (a) Upon selection, an agent learns the `party` value of each other agent in their circle; denote these by $\texttt{party}_i$, for $i = 1, 2, \ldots, n$.

   (b) The agent calculates a weight, w_i, for each of their neighbor interactions based upon the strength of their party affiliation. Specifically,

   $$w_i = \begin{cases} \texttt{party}_i, & \text{if } \texttt{party}_i \geq 0.5 \\ 1 - \texttt{party}_i, & \text{if } \texttt{party}_i < 0.5 \end{cases},$$

   for $i = 1 \ldots n$. This calculation ensures that interactions with both blue and red agents are treated identically with respect to weight. Moreover, neighboring agents with "extreme" `party` values—those close to 0 or 1—are given higher weights than neighboring agents with moderate `party` values close to 0.5.

   (c) The selected agent then computes a `similarity` score, by adding the weights of all agent interactions from their own party and dividing by the total of the weights for all interactions. Let Ω denote the set of other agent numbers for whom $\texttt{party-cat}_i$ matches the `party-cat` of the querying agent. That is,

   $$\texttt{similarity} = \frac{\sum\limits_{i \in \Omega} w_i}{\sum\limits_{i=1}^{n} w_i}.$$

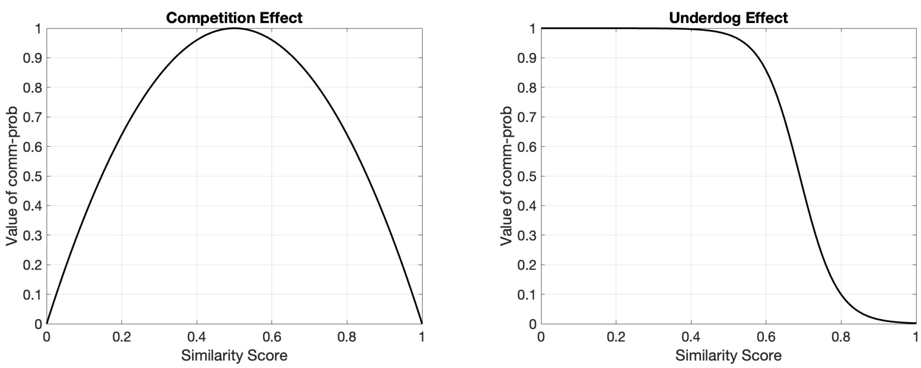
   (d) The agent then uses their computed `similarity` score to determine their `comm-prob` value, or their probability of participating in the election based on community factors. The computation method for this value is determined by the user's choice of the environment `comm-effect` parameter, which can be set to either "Competition Effect" or "Underdog Effect" (both of which were discussed in Section 2.1.4). For each of these community effects, we define a function that takes the `similarity` score as an input and outputs a corresponding probability of voting based on community factors. For the Competition Effect, our function is

   $$\texttt{comm-prob} = -4(\texttt{similarity} - 0.5)^2 + 1,$$

   which is shown in Figure 3(a). We defined this function such that voters who have a similarity score close to 0.5 (indicating roughly equal interaction with members of both parties) are most incentivized to vote, while those perceiving themselves as firmly in the majority or minority based on community interaction become apathetic. Parameters in the equation were chosen to mimic this phenomenon empirically. For the Underdog Effect, we implement the reverse sigmoid function

$$\texttt{comm-prob} = 1 - \frac{1}{1 + 6e^{-20(\texttt{similarity}-0.6)}},$$

shown in Figure 3(b). When determining the function's parameters, we wanted to ensure that a similarity score of 0.5 still corresponded to a `comm-prob` close to 1. This supports the idea that in this scenario, agents who perceive themselves to be in the minority are highly incentivized to vote, while agents interacting mainly with other like-minded agents are apathetic. The parameters used to generate the community probability graphs in Figure 3 are chosen to mimic the described effects from Levine and Palfrey (2007) [26], and were not varied in any subsequent analysis. We leave the study of additional variations of these two functions for future work.



(a) (b)

Figure 3: Community-based probability of voting as a function of similarity score under (a) the Competition Effect, as compared to (b) the Underdog Effect.

Procedure for `elect` Submodel:

1. Agents are selected in a random order. For each agent:

   (a) The agent computes their `vote-prob`, or overall probability of voting in the election, by combining their `demo-prob` and `comm-prob` values in a weighted linear combination using the environment parameter `a`. That is,

   $$\texttt{vote-prob} = \texttt{a} \cdot \texttt{demo-prob} + (1 - \texttt{a}) \cdot \texttt{comm-prob}.$$

   With such a weighting, an `a` value of 1 allows a user to study what would happen if agents made their decision to vote based solely on their individual characteristics. Conversely, setting `a` $= 0$ allows a user to study what would happen if voters were entirely swayed by their interactions with their community members.

(b) An agent "makes a decision" whether to vote or not by comparing their `vote-prob` value to a randomly-generated number on the interval $[0, 1]$.

i. If their `vote-prob` exceeds the generated random number, they will cast their vote in the election by incrementing the value of `count-total`—a value documenting the number of cast votes in an election run, ranging from 0 (reflecting an entire population of abstainees) to 2,601 (representing a voter turnout rate of 100%)—as well as their party-specific vote total of `count-blue` or `count-red`.

ii. Else, if their `vote-prob` does not exceed the chosen generated random number, the `pcolor` value for their corresponding grid cell is set to black to indicate their decision to abstain on the user interface.

2. Finally, the number of counts for the winning party, as well as the overall voter turnout percentage (which is calculated by dividing `count-total` by 2,601, the number of "registered voters" in our simulation), are displayed on the user interface.

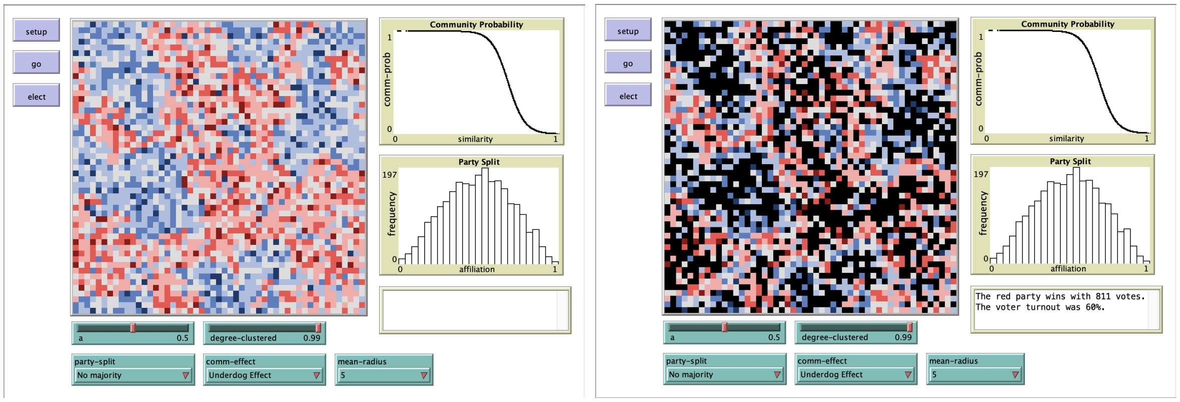
This concludes our Overview, Design concepts, and Details (ODD) protocol.

2.2 User Interface

The model described above was implemented in NetLogo 6.3.0 [32]; all code is provided at: `https://github.com/gbrophy2/abm-voter-behavior.git`. The starting user interface is shown in Figure 4(a). In the main view, each agent's political affiliation is indicated by a colored patch chosen along a spectrum, with dark red representing "strongly Republican," white representing "moderate," and dark blue representing "strongly Democrat." Toggles are provided for the five varying environment settings:

- `a`: Varies from 0 to 1 in increments of 0.1
- `degree-clustered`: Varies from 0.5 to 1 in increments of 0.01
- `mean-radius`: Selected from 4, 5, or 6
- `comm-effect`: Selected from "Competition Effect" or "Underdog Effect"
- `party-split`: Selected from "Strong Red Majority," "Partial Red Majority," "No Majority," "Partial Blue Majority," or "Strong Blue Majority"

To the right of the agent view, graphs indicating the distributions of community probabilities and party affiliations are provided for the given scenario; in Figure 4(a), we show a sample setup for the choices `a` $= 0.5$, `degree-clustered` $= 0.99$, `party-split` = "No Majority," `comm-effect` = "Underdog Effect," and `mean-radius` $= 5$.



(a) (b)

Figure 4: (a) Initial user interface with model settings `party-split` = "No Majority," `a` = 0.5, `degree-clustered` = 0.99, `comm-effect` = "Underdog Effect," `mean-radius` = 5, (b) The same user interface after running the `election` command, where abstaining agents are denoted by black squares.

Figure 4(b) shows the user interface with the initialization from (a), once the simulation has been run. Grid cells—or patches—for agents that abstained from the election have been colored black to indicate their non-participation. Additionally, the final results for the winning party and overall voter turnout are displayed on the bottom right. In this scenario, the victory went to the red party, with an overall voter turnout of 60%.

2.3 Sensitivity Analysis Methodology

Our model output relies on five input settings and/or parameters: two categorical inputs (`party-split` and `comm-effect`), and three numerical inputs (`mean-radius`, `a` and `degree-clustered`). Because this model has not been calibrated to existing data, values for each of these settings that would map to real election results remain unknown. In our presentation of results, we focus on quantifying how the model output relies on parameter combinations chosen from across the permitted ranges. Specifically, we consider the range and distribution of model outputs across a full sweep of the input space. For this parameter sweep, we vary parameter `a` from 0 to 1 in increments of 0.1 and `degree-clustered` from 0.5 to 1 in increments of 0.1. The `party-split` setting is investigated for all five categories: "Strong Red Majority," "Partial Red Majority," "No Majority," "Partial Blue Majority," and "Strong Blue Majority." The parameter `mean-radius` is permitted values of 4, 5, and 6, while `comm-effect` is toggled between "Competition Effect" or "Underdog Effect." Each of the 1980 resulting combinations is replicated 50 times using the BehaviorSpace tool in NetLogo [33] and then averaged to control for stochasticity. With model inputs and corresponding outputs from each of these runs, we can then quantify how sensitive the model is to perturbations in the input settings using a sensitivity analysis.

As detailed in [27], a robust sensitivity analysis is often omitted in agent-based modeling, due to the computational expense required to perform even a simple parameter sweep, and the tendency of ABMs to rely not only on input parameters but also on non-parametric elements. However, it is imperative to conduct such analyses in order to draw meaningful conclusions from the model. Through this work, we target the four goals of sensitivity analysis discussed in Borgonovo et. al (2022) [27]: factor prioritization, parameter interactions, robustness, and direction of change. Factor prioritization involves determining which of our inputs contribute most to changes in our outputs. Similarly, sensitivity analysis techniques targeting interaction analysis explore which combinations of inputs yield the greatest impact on our output quantities when acting in combination. Tools focusing on robustness examine the consistency of these impacts. Finally, analyzing direction of change determines whether perturbations in input parameters yield output changes that are positive or negative. As outlined in Figure 5, we analyze the impacts of inputs `a`, `degree-clustered`, `comm-effect`, `party-split`, and `mean-radius` on two different quantities of interest (QoIs): the overall voter turnout percentage, detailed in Section 3.2.1, and the extent to which the party breakdown of the final election result differs from the initial electorate split, discussed in Section 3.2.2. For the latter, we specifically investigate circumstances under which this discrepancy leads to a minority party victory. Such analysis re-contextualizes our turnout results to highlight which conditions create the largest discrepancies between original simulation `party-split` and the margin of victory for the winning party.

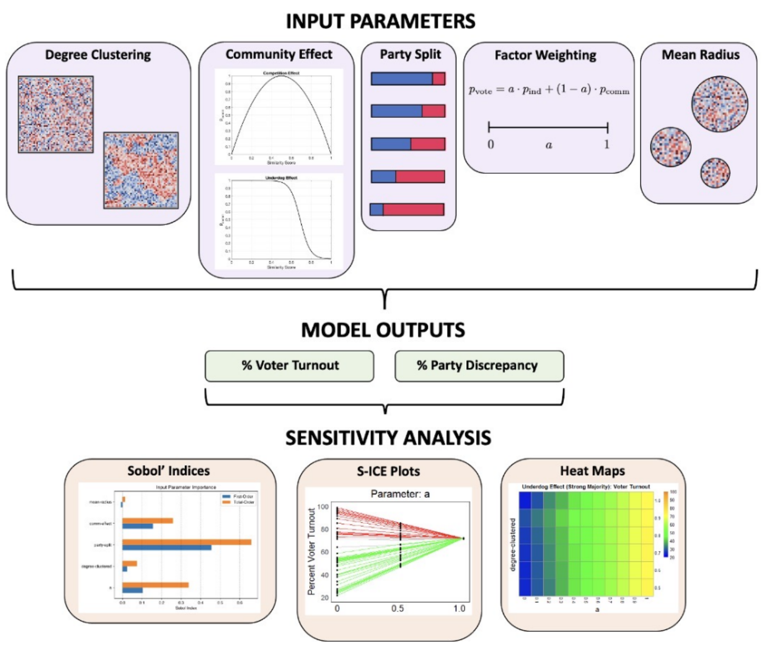


Figure 5: Schematic detailing the model analysis phase of the investigation.

To investigate our four goals of factor prioritization, interactions, robustness, and direction of change, we use a combination of sensitivity analysis tools and plots: stochastic-individual conditional expectation (S-ICE) plots, Sobol' indices, and heatmaps detailing effects of interacting parameters, each of which is briefly described here.

Stochastic-Individual Conditional Expectation Analysis:

One-at-a-time sensitivity analysis is used to observe the individual effects of each input on the quantity of interest, allowing us to qualitatively examine the robustness of the model and the direction of change in model output with respect to each parameter. Here, we use stochastic-individual conditional expectation (S-ICE) plots to examine the effect of the input parameters on the quantity of interest. S-ICE plots are created by varying each parameter across its permitted range while holding all other parameters constant [27]. For our model analysis, we vary `comm-effect`, `party-split`, and `mean-radius` across each of their possible categorizations. The parameter `a` is examined at discrete values 0, 0.5, and 1.0, to observe the impact of heavily weighting either `comm-prob` or `demo-prob`, or equally weighting them. Likewise, `degree-clustered` is discretized to values of 0.6, 0.8, and 1.0 to observe the impact of very little to extreme clustering of agents on voter turnout. Each pair of points represent scenarios in the ABM that differ only in the value of the perturbed parameter—to control for stochasticity, each point in our S-ICE plots represents the aggregated mean of a set of 50 model replicates for the given parameter set. The sign of the line's slope indicates the direction of change in the model output as the parameter is varied. The statistical significance of each test is determined using a two-tailed Welch's t-test with a significance level of $\alpha = 0.05$.

Sobol' Indices:

Global sensitivity analysis allows us to consider the sensitivity of the model output to a large range of possible parameter combinations, an ideal analysis in scenarios where the values of the input settings are uncertain. Sobol' indices are a commonly used variance-based tool for global sensitivity analysis; they decompose the variability in the model output and attribute it to specific inputs or combinations of inputs. Additionally, the Sobol' indices reveal the ranking—or factor prioritization—of the inputs in order of importance to the model output.

The calculation of first-order Sobol' indices indicates which inputs have the greatest contribution to the change in model output on their own. Specifically, the first-order Sobol' index for a given input quantifies how much of the variation in the model output is due to varying that input alone, averaged over variations of all other inputs. Total-order Sobol' indices can also be calculated to quantify the effects of a given input when acting in combination with all other inputs; in essence, an input's total Sobol' index is defined as the sum of all sensitivity indices that involve that input, from first-order to mth-order, where m represents the total number of inputs. A full treatment of Sobol' indices and their mathematical derivation can be found in [34, 35].

For this analysis, we utilize the sensitivity analysis library in Python, SALib, and its accompanying Sobol' indices functionality [36]. Specifically, we instantiated our NetLogo model in Python using the pyNetLogo package [37]. Then, we used SALib to generate a Sobol' sample from our input space. We ran the model for each combination of parameters from the sample using pyNetLogo, then calculated first and total-order Sobol' indices using SALib. For our model setup, Sobol' first-order and total-order index estimates were found to converge for $n \geq 500$; we use a sample size of $n = 1,024$ for all figures presented in our analysis.

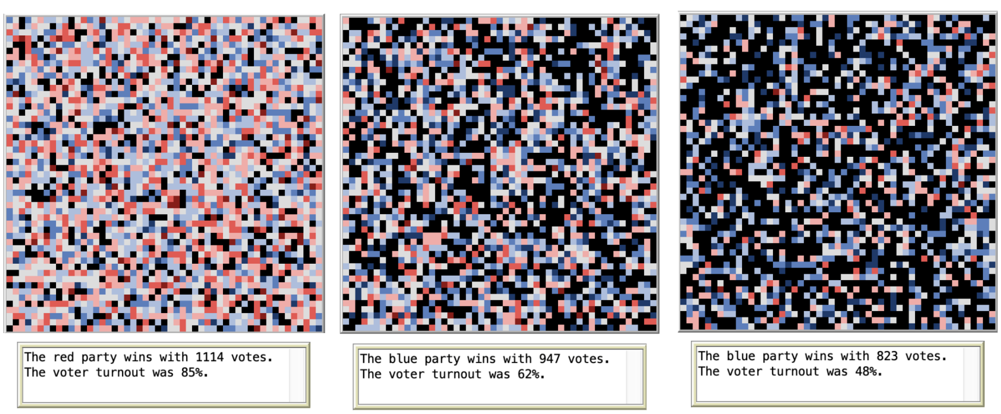
Interaction-Based Heatmaps:

To analyze how two inputs may act in combination to produce particular outcomes in the quantity of interest, we also include two-at-a-time interaction heatmaps. In these figures, two quantitative parameters are varied across a discretized space, with all other inputs held fixed. In particular, our focus is on varying `a` on the range $[0, 1]$ and `degree-clustered` on the range $[0.5, 1]$, both in increments of 0.1, in tandem, for fixed combinations of the categorical and discrete variables `comm-effect` and `party-split`.

3 Results

3.1 Individual Case Studies

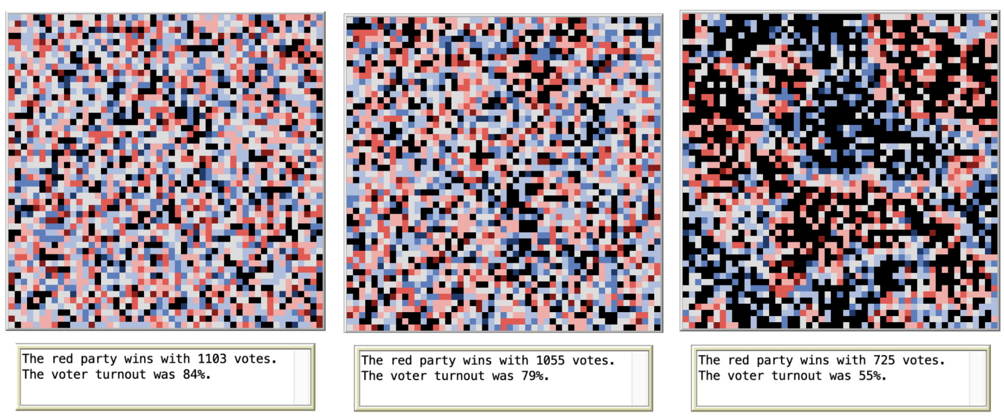
To begin, we present a few single-input case studies to illustrate how changing just a single setting in the model can strongly impact the overall voter turnout. For instance, in Figure 6, we display the final results for a scenario in which only the `party-split` input is varied (from "No Majority" on the left, to "Partial Blue Majority" in the middle, to "Strong Blue Majority" on the right). All other inputs are held constant. When there is a good mix of voters from both parties and no clustering, as on the left, all voters sense a competitive atmosphere and are highly incentivized to vote, yielding an overall turnout of 85%. On the far right, however, when the majority of voters are from the blue party, they become apathetic in response to being surrounded by like-minded people. In this scenario, only 48% of voters choose to participate.



(a) (b) (c)

Figure 6: Varying `party-split`. (a) Final user interface with `party-split` = "No Majority," (b) with `party-split` = "Partial Blue Majority," (c) with `party-split` = "Strong Blue Majority."

Figure 7 illustrates how the value of the `degree-clustered` variable impacts the arrangement of the agents on the grid and their resulting choices about participation. With `degree-clustered` $= 0.5$, the choice of each agent's party affiliation is entirely random. As we move towards the upper end of the spectrum (shown for `degree-clustered` $= 0.99$ on the far right), the agents increasingly cluster with other like-minded agents. Because of this clustering, agents perceive themselves to be in the majority party, even though the party affiliations are actually balanced in number across the entire view. Due to their mistaken perception, many agents choose not to participate in the election, and overall voter turnout declines from 84% with `degree-clustered` $= 0.5$ to 55% with `degree-clustered` $= 0.99$. Variations of both `degree-clustered` and `party-split` can greatly alter the variety of similarity scores present in the electorate. While all scores fall within the range of $[0, 1]$, across a full parameter sweep varying all inputs we observe a mean maximum similarity score of 0.65 and a mean minimum similarity score of 0.07. These cases correspond to electorates with more socio-political homogeneity, like Figure 6(c), and electorates with higher overall socio-political heterogeneity, like Figure 6(a), respectively.



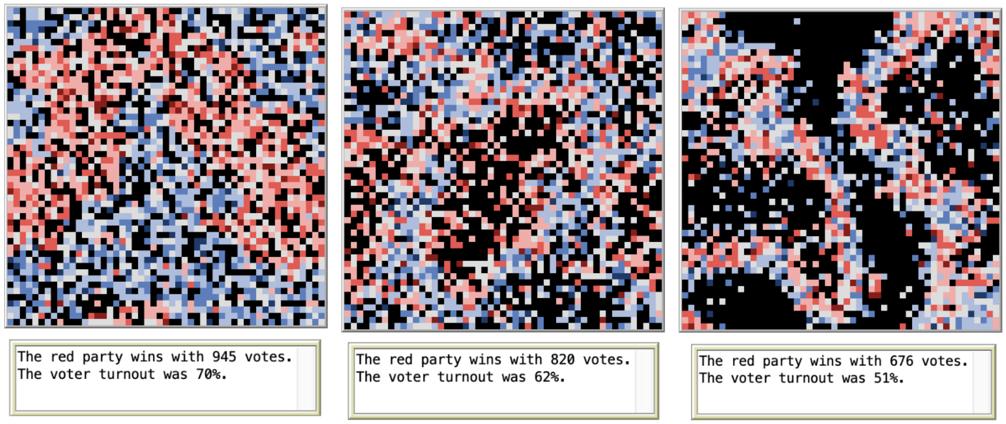
(a) (b) (c)

Figure 7: Varying `degree-clustered`. (a) Final user interface with `degree-clustered` $= 0.5$, (b) with `degree-clustered` $= 0.75$, (c) with `degree-clustered` $= 1.0$.

Finally, in Figure 8, we illustrate how the weight chosen for the linear combination of the individual and community-based probabilities affects the final result, with the `degree-clustered` variable set to its highest value of 1.0. With `a` $= 1.0$ (i.e., voters make their decisions based entirely on individual, demographic-based choice), there is a relatively high voter turnout. As we decrease `a`, the voter turnout decreases, since voters now make their decisions based on their surrounding community, and many elect not to vote as a result of being surrounded by like-minded people.

3.2 Sensitivity Analysis

Next, we present the results of our sensitivity analysis, using the plots and metrics discussed in Section 2.3. First, we consider overall voter turnout as our quantity of interest



(a) (b) (c)

Figure 8: Varying `a`. (a) Final user interface with `a` = 1.0, (b) with `a` = 0.5, (c) with `a` = 0.0.

(Section 3.2.1), and then turn our attention to investigating which inputs are most likely to cause a major discrepancy between party affiliation in the electorate and party affiliation in actual votes cast (Section 3.2.2).

3.2.1 Voter Turnout QoI

S-ICE plots are presented for each of the five inputs in Figure 9, with regard to their impact on overall voter turnout. In general, we observe a statistically significant downward trend in voter turnout when switching from "Competition Effect" to "Underdog Effect," while keeping all other settings constant. This reflects a tendency for majority party voters to be more apathetic in the Underdog scenario, an outcome that was observed in Table 2. Similarly, we observe a downward trend in voter turnout when we move from "No Majority" to "Partial Majority" to "Strong Majority" party splits; the increasing number of majority party voters yields the perception of a less competitive scenario, increasing apathy among majority party voters. The trend in the `degree-clustered` plot is also predominantly decreasing; as the tendency of voters to communicate in echo chambers increases, voters are more likely to perceive themselves to be surrounded primarily by members of their own party, leading to voter apathy among both parties in response to the perception that a vote will not meaningfully impact the election outcome.

In contrast, the trend for `mean-radius` indicates that voter turnout tends to increase as the average number of interactions among agents increases, due to the fact that increasing the number of interactions among agents helps each agent to overcome the effects of any imposed clustering. A larger number of interactions makes it more likely that an agent will observe members of both parties within their social circle, and adjust their likelihood of voting upward in response to this perceived level of competition.

Among all five inputs, parameter `a` is the only parameter for which we observe that the direction of change may regularly be either positive *or* negative, depending on the other settings in the model. Recall, the value `a` = 1 indicates that participation in the election depends entirely on a voter's demographic characteristics, as dictated by their `demo-prob` values that were estimated from the 2016 CCES study [29]. Thus, for `a` = 1, all aggregated means are clustered around the mean value produced from the logistic regression model, since each of the other community-based factors (`degree-clustered`, `comm-effect`, `mean-radius`, and `party-split`) have no impact on the results. For small values of `a`—where the focus is almost entirely on community impact—these community-based factors have a large impact on overall voter turnout. For input combinations yielding a low voter turnout for `a` = 0, we observe an increase in voter turnout as we hold all other parameters constant but shift `a` from 0 to 0.5 to 1. On the contrary, input combinations yielding a voter input larger than the demographic-based average when `a` = 0 display a downward trend in voter turnout as we increase `a`. This indicates that varying `a` has a large impact on the *variability* of voter turnout outcomes, but little impact on the *average* voter turnout over many simulations.

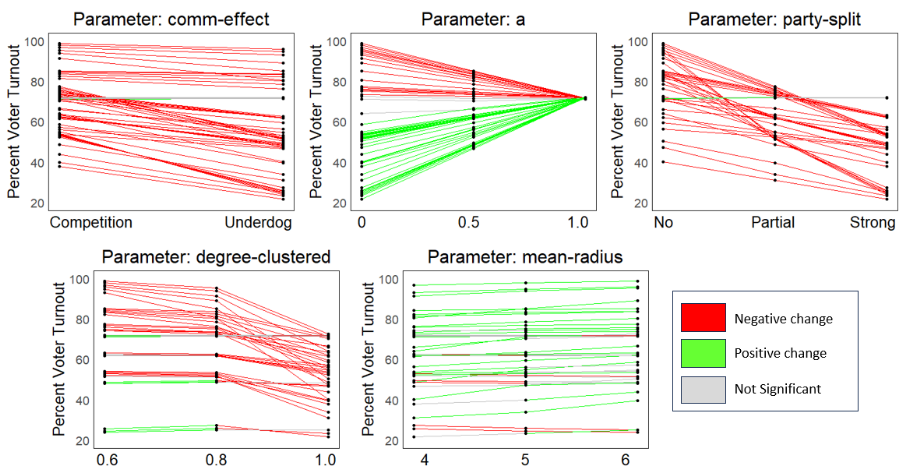


Figure 9: Stochastic-individual conditional expectation (S-ICE) plots for the five input parameters.

Figure 10 displays the first and total-order Sobol' indices of our input settings with regard to the voter turnout QoI. The results of this analysis can be used to fulfill our goals of factor prioritization as well as understanding interactions between various inputs. The `party-split` parameter has the highest impact on voter turnout both individually and in conjunction with the other inputs. On its own, `party-split` characterizes the number of members of each party, impacting how agents view their community based on this distribution alone. Voters residing within a community with a strong majority presence are either disincentivized to vote due to the lack of competition in their environment, or perhaps incentivized to participate if they are a member of the minority party operating under the Underdog Effect. In combination with inputs such as `a` and `comm-effect`, `party-split` has an even larger effect on voter turnout. Depending on which community effect is set and how heavily it influences agents' decisions to vote, the ratio of agents affiliated with each party is very impactful.

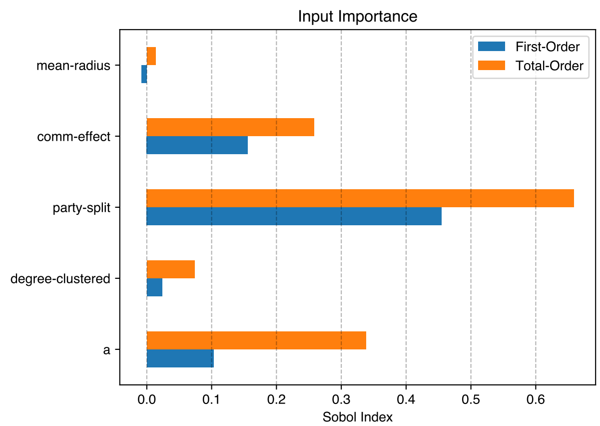


Figure 10: First and total-order Sobol' indices for voter turnout quantity of interest.

The `comm-effect` setting is the second most influential input when considered on its own; however, parameter `a` is highly influential when considered in conjunction with other settings. For instance, scaling `a` from 0 (where agents make their decisions entirely based on community factors) to 1 (where agents act on an individual basis) has the effect of dialing the impacts of other community-based factors such as `comm-effect`, `degree-clustered` and `party-split` down.

Finally, we observe that the `mean-radius` parameter has a very minimal impact on voter turnout, reflected in both its first and total-order Sobol' indices. This indicates that varying `mean-radius` has little impact on our quantity of interest, allowing us to fix the value of this parameter in later analyses to improve computational efficiency without sacrificing a robust investigation.

Motivated by the results of our Sobol' analysis, we now turn to investigating how parameters and settings may act in combination to produce varied outcomes, analyzing both input interactions and direction of change. Figure 11 showcases the results for both of the Competition and Underdog Effects across each of the "No Majority," "Partial Blue Majority," and "Strong Blue Majority" cases. Without loss of generality, we do not include plots for "Partial Red" and "Strong Red" majorities here, since the results of voter turnout are essentially identical to the blue party majority results. In each heatmap, we vary the numerical parameters `a` and `degree-clustered` to quantify how they interact to produce different voter turnout outcomes. Note that since `mean-radius` was not found to be strongly influential with regard to overall voter turnout, we fix `mean-radius` at 5 for this heatmap analysis.

As we move from left ("No Majority") to right ("Strong Blue Majority") under the Competition Effect in Figure 11(a), we see that overall voter turnout decreases. When party affiliations are well-balanced, voters sense a competitive atmosphere and are driven to the polls. In contrast, when one party strongly outnumbers the other, members of both parties become apathetic, sensing that their vote will have little impact on the election outcome. We also observe a decrease in voter turnout as `degree-clustered` increases. As agents increase their tendency to aggregate in echo chambers with agents of similar

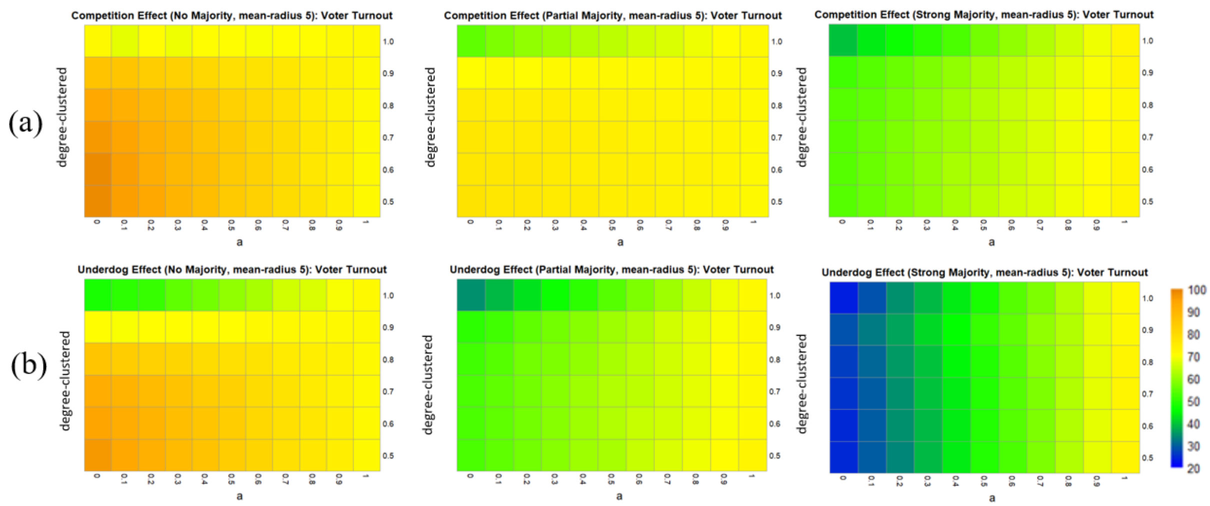


Figure 11: A comparison of the voter turnout QoI for the (a) Competition and (b) Underdog Effects with `mean-radius` $= 5$, under varying conditions of `a`, `degree-clustered`, and `party-split`.

party affiliation, they perceive less competition around them, becoming more apathetic in alignment with the imposed Competition Effect. Additionally, we observe that this dependency of voter turnout on `degree-clustered` is more impactful for low values of `a`, when agents are putting a greater weight on their community interactions when choosing whether to participate in the election. These results reinforce what was found in the Sobol' analysis: parameters `a` and `degree-clustered` are more impactful when considered in combination, rather than in isolation.

Figure 11(b) presents the results of the parameter interactions under the Underdog Effect. In contrast to the Competition Effect plots, we see an overall lower voter turnout across all three `party-split` scenarios, and especially so for the two scenarios in which one party holds a clear majority. In these scenarios, voters in the majority party are disincentivized to vote based on their community interactions. Agents in the minority party are activated, but represent a much smaller portion of the vote due to their minority status, resulting in a lower overall turnout. The parameters `a` and `degree-clustered` have similar impacts on the quantity of interest as in the Competition Effect plots.

It is important to note that because our model simulates an arbitrary electorate (that is, it is not calibrated to conform to a specific data set), we do not expect every model outcome to map directly to an observed data point. Rather, our goal is to understand the full distribution of model outcomes that can be produced over a wide range of possible input combinations and to ensure that our model is capable of producing the range of outcomes that we observe in reality. To do so, we examine the distribution of voter turnouts produced by our ABM and compare it to data collected from the 2016 U.S. general election and the 2018 U.S. midterm election cycles. For the 2016 general election, we report state-level voter turnout statistics, where voter turnout is defined as the percentage of total votes out of the eligible voting population for all 50 U.S. states and the District of Columbia [38]. This data set gives us a rough idea of how voter turnout rates may vary as we compare states with a strong preference for one candidate over the other to swing states with close vote counts for each. However, because our ABM is built to look at voting patterns on a much smaller scale (analyzing several thousand voters at a time), we also include county level data from Pennsylvania [39, 40], a key swing state that includes heavily-conservative counties (e.g., Armstrong, Bradford, and Clearfield Counties), liberal-leaning urban counties (e.g.. Allegheny and Philadelphia Counties), and moderate counties that are often split nearly 50-50 in their votes cast (e.g., Northampton County), for a better comparison based on population size. This county-level data is then reported again for the 2018 midterm elections, to allow for any juxtaposition between elections in which the presidency is at stake versus elections that are more local in nature, two scenarios that tend to exhibit different voting trends. Numerical summaries for each data set and the ABM outcomes are reported in Table 2.

Data	**Min**	**Q1**	**Median**	**Mean**	**Q3**	**Max**
2016 U.S. States	47.23	59.30	62.50	62.44	66.11	74.22
2016 PA Counties	61.16	68.77	71.77	71.04	73.95	77.77
2018 PA Counties	47.57	56.42	59.68	58.62	60.92	66.69
Model	**Min**	**Q1**	**Median**	**Mean**	**Q3**	**Max**
ABM (All Scenarios)	21.86	56.83	67.38	64.58	73.46	99.05
ABM (Competition)	37.85	63.54	71.86	69.89	74.89	99.05
ABM (Underdog)	21.86	50.13	61.23	59.27	70.07	96.85

Table 2: Comparison statistics between voter turnout in 2016 U.S. states and simulated voter turnout from the ABM.

While it can be observed from Table 2 that our ABM outcomes exhibit larger ranges in turnout than the state- or county-level data, we note that it may be the case that some of the model settings used to generate these extreme outcomes are in fact not observed in reality. For instance, we can see from Figure 11(b) that the only place in which we see turnout in the 20-30% range for the ABM is for scenarios in which there is a strong majority present, voters behave in accordance with the Underdog Effect, and our `a` parameter is close to 0. In practice, this would mean that voters are voting almost entirely based on community-driven factors, and all of the members of that strong majority abstain from voting because of the Underdog Effect. This is not likely to happen in reality, and so the fact that this outcome seems extreme compared to others is not troubling. Likewise, we only observe ABM voter turnout outcomes close to 100% when `a` is close to 0, `degree-clustered` is set to 0.5, and there is no majority; this represents a perfectly mixed population in which everyone is driven to vote—regardless of whether the Underdog or Competition Effects are imposed—because everyone senses themselves to be in a 50-50 position and chooses to vote based solely on their community-based voting probability. As with our previous example, this scenario is unlikely to occur in reality; members of the electorate are likely to use some combination of individual drive and community factors to make their choice, and a perfectly mixed population is unlikely to appear. With these extreme cases not occurring in reality, we focus on the fact that the center of our ABM outcome distribution (focusing between the first and third quartiles) is able to produce the full range of outcomes that we observe in real life. Further calibration of our model parameters to specific electorates would allow us to determine realistically feasible values for each of our parameter settings.

Patterns emerging from our sensitivity analysis may also be used to explain trends that are occurring on an individual basis. For instance, the tendency of the Competition Effect to produce higher overall voter turnout rates—as reflected in both Figure 11 and Table 2—is reflected in county level data from the 2016 U.S. presidential election. For instance, Suffolk County, New York had the largest surge in voter turnout—an increase of 6.87%—among all 206 "pivot counties": counties who voted for Democratic candidate Barack Obama in 2008 and 2012 but for Donald Trump, the Republican nominee, in 2016. According to a November 2016 voter registration report, Suffolk County had roughly the same amount of registered Republicans and Democrats (about 300,000 each) before the 2016 presidential election [41]. Sensing this tight race—and in turn, mimicking the Competition Effect—voters from both parties were highly incentivized to participate in the election, driving up overall turnout to 64.46%.

3.2.2 Discrepancy in Party Proportion QoI

We now consider another quantity of interest which tracks the change from the initial percentage of a party in the electorate—or total agent population—to that party's share of the votes in the election. We subsequently refer to this output as the "discrepancy in party proportion." For example, a shift in the red party from occupying 30% of the electorate to 55% of the votes in the election leads to a discrepancy in party proportion of +0.25. The corresponding shift in the blue party is then 70% to 45%, resulting in a −0.25 change. Because these shifts are equal and opposite, we can follow the change in composition of both parties throughout model simulations by measuring the change in one. Considering this output allows us to target "upset" cases where the minority party wins the election due to substantial apathy among majority party voters.

We first repeat our Sobol' analysis using this new quantity of interest. The results of this analysis are presented in Figure 12. While `party-split` is the only input with a significant first-order Sobol' index, we observe that it is in input *interactions* where most of our model output sensitivity resides. When considering the total-order Sobol' indices, `party-split`, `comm-effect`, and `a` emerge as the most influential inputs. This result is unsurprising, as the combination of `party-split` and `comm-effect` dictates how voters are motivated in response to the overall ratios of party affiliations in their community, with `a` dictating how much weight they place on this perception when making their decision to vote. Once again, we observe that `mean-radius`—as well as `degree-clustered`—are relatively uninfluential. We once again fix `mean-radius` at 5 for the remainder of this analysis.

Finally, we present heatmaps detailing how the remaining four inputs interact to produce various outcomes. Without loss of generality, we present the results for the "No Ma-

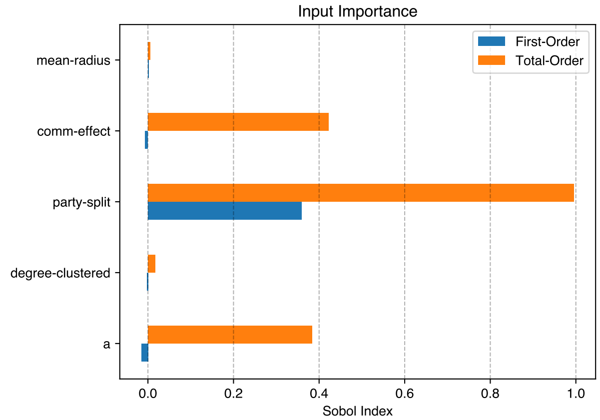


Figure 12: First and total-order Sobol' indices for party discrepancy quantity of interest.

jority," "Partial Blue Majority," and "Strong Blue Majority" scenarios, with the quantity of interest being the discrepancy in party proportion for the *red* party.

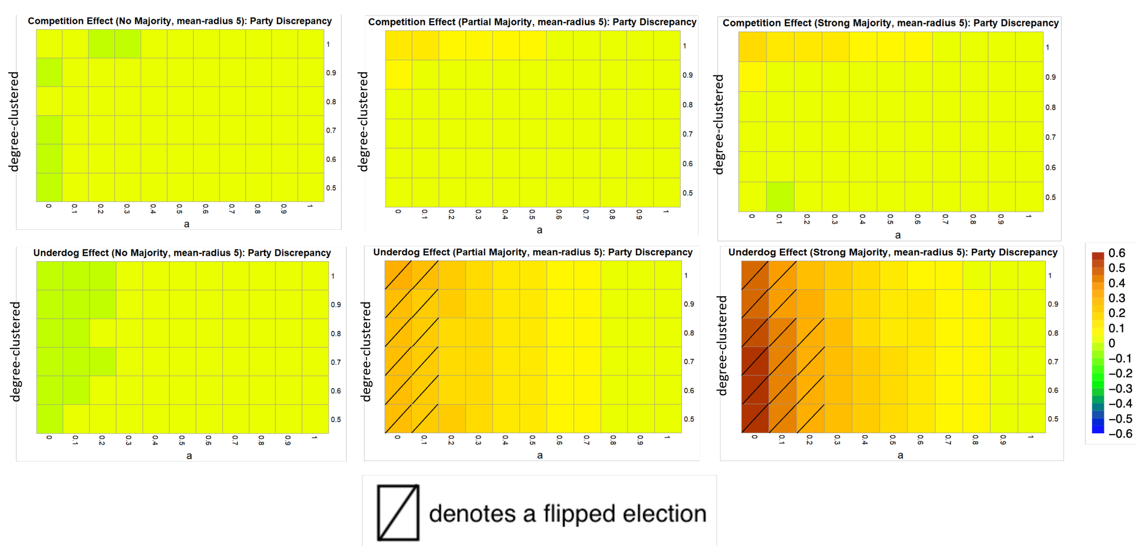


Figure 13: A comparison of the party discrepancy QoI for the (a) Competition and (b) Underdog Effects with `mean-radius` $= 5$, under varying conditions of `a`,`degree-clustered`, and `party-split`.

Figure 13 shows the discrepancy in party proportion for the Competition Effect (top) and the Underdog Effect (bottom), across the three `party-split` choices. Cells with strikethroughs represent scenarios in which the minority candidate wins the election; that is, the discrepancy in party proportion is large enough to tip the election in favor of the candidate with whom fewer members of the electorate favor. It is clear that within the Competition Effect scenarios, the discrepancy in party proportion QoI is relatively insensitive to `degree-clustered`, `a`, and `party-split`. However, for the Underdog Effect, we observe the potential for extreme discrepancy in scenarios in which there is a clear majority for one party. In particular, for small values of `a`—indicating that voters place a high weight on their community interactions—and categories of either "Partial Majority" or "Strong Majority" for `party-split`, the outcome of the election is likely to favor the minority candidate, due to extreme apathy among majority party voters.

Such a "flipped" election outcome has been observed in reality in regions that mimic the Underdog Effect. For example, in 2016, Lee County, Iowa experienced 58.76% voter turnout in the 2016 presidential election [42]. At a rate 9.98% lower than their 2012 presidential voter turnout, Lee County saw the highest decline in voter turnout from 2012 to 2016 of all 206 pivot counties (that is, counties that voted for Barack Obama in 2008 and 2012, and then voted for Donald Trump in 2016). Voter registration reports from November 2016 show that Lee County had a roughly 2:1 ratio of registered Democrats to registered Republicans [43]; however, only 39% of electoral ballots selected the Democratic nominee for president compared to 55% for the Republican candidate [44]. Our model's ability to reproduce such outcomes—in which an area with a strong Democratic majority votes in favor of the Republican candidate—is further evidence of the flexibility of our framework.

4 Conclusion

In this work, we have presented an agent-based model developed to understand how an eligible voter's perception of the political climate in their community may impact their decision to participate in an election. By controlling for individualistic characteristics that may motivate voters using data from the 2016 CCES survey [29] while allowing various community-based factors to vary (e.g., the degree to which voters cluster by party affiliation, the initial party affiliation breakdown in the community, the level of activism within the minority party, etc.), we determine the conditions under which voter turnout would surge or decline, as well as identify combinations of settings that are likely to lead to an election with an outcome that differs from the majority opinion in the electorate. We remark that this model is intended to be used on a generalized scale; it was not calibrated to match data from a particular region, and care should be taken when drawing conclusions. Specifically, we considered the cost of model complexity as a major factor—the goal was to focus on a few settings and understand which inputs may play the largest role in driving voter apathy. While many other factors (or other levels of the existing factors) could be incorporated, these may come at the cost of obfuscating the major driving mechanisms of the model output. Thus, this model serves as a first step towards a deeper investigation into community-based drivers of voter apathy.

Through a robust sensitivity analysis, we discovered that our quantities of interest were highly influenced by the initial breakdown in party affiliation among eligible voters in the community, as well as the type of motivating community effect imposed ("Competition" versus "Underdog,") and the degree to which voters weight their community interactions against their individualistic characteristics when making their decision to vote or abstain. In particular, these three factors are important not just in isolation, but even more so when considered in combination. For instance, we find that in order for a "flipped" election to occur within our modeling framework—that is, for the election winner to be the minority candidate—we must observe a scenario in which voters in the minority party are very active ("Underdog Effect"), the party affiliation breakdown is highly skewed toward one party over the other, and eligible voters place a large emphasis on their interactions with those in their community when determining their inclination to vote.

While our model was developed as a first step towards investigating which combinations of community-based factors might be most influential in determining the level of voter apathy in a community, much more could now be done to place this in a specific context. Selecting a specific region and calibrating the model settings to match data and insights from that region would allow for more robust model validation and enable meaningful insight into the major driving mechanisms of voter apathy in that region. For instance, it may be the case that in a given community, one party is more "vulnerable" to community interactions—in the sense that they are more likely to be swayed to vote or abstain based on their interactions with those around them—while the other party may contain individuals that are more likely to make their decision based on their own core values. This type of discrepancy could be accounted for by allowing the `a` parameter, which serves as the weight in the linear combination combining demographic and community factors, to vary among parties or even different demographic characteristics such as age or education. Similarly, knowing the political history of a particular community may give insight into past patterns of mobilization of the minority party, which could allow a user to better understand what kind of community effect should be imposed; as with parameter `a`, this community effect could also be altered on an party-wide or even individualized basis. Additionally, while the `party` variable is currently modeled as a continuous variable on the interval $[0, 1]$, this is a simplification that assumes a member of the electorate lands at the same location on the political spectrum for all political issues. In reality, it is often the case that voters have more complex and multidimensional political stances; for instance, a voter may consider themselves conservative with respect to fiscal matters but lean more liberal when it comes to social policies. In future iterations of the model, the `party` variable could be adapted to model this multidimensional political framework, particularly if researchers were interested in studying multiple genres of political issues simultaneously. Finally, another factor driving voting behavior that our model could incorporate in the future is political polarization. To account for a growing ideological divide in the current political climate in our model, one could introduce an additional option for `party-split` where agents' `party` values are sampled from a bimodal distribution skewed toward more "extreme" opinions. While our model as it currently stands is intended to be very generalized in order to allow for the determination of large-scale effects in a simplistic environment, further adaptations such as those discussed here would be necessary to ground the model in a particular regional context.

Additionally, we note that our model was developed specifically to target questions about community influence of voter apathy. However, this model could also be adapted to explore other related questions of interest. For example, an interesting next step might be to allow for eligible voters to change not only their decision about whether to vote or abstain, but to actually change their party affiliation based on interactions with those in their community. Indeed, recent research suggests that extreme voters may be more likely to become more moderate over time, as opposed to moderate voters becoming more extreme in their preferences. An extension of this model to incorporate multiple time steps in which voters with extreme stances gradually drift towards the moderate viewpoint could make the simulation more realistic [45]. Additionally, while the model interface currently assumes that all of the agents in the model are participating in the same election, including district lines within the interface—such that voters are casting their votes for different sets of candidates while still possibly interacting with voters from the other side of the divide—might provide an interesting look into how gerrymandering plays a role in voter apathy. Larger abstractions from our current framework, potentially through the use of differing demographic datasets, could include redefining the political structure to analyze apathy in multiparty systems, culturally extending our basis away from the United States. Further model analysis of the impact of factors like social or mainstream media, grassroots campaigns, and voter suppression tactics could also help contextualize the current political climate. An investigation into any of these adaptations and their ensuing effects on voter apathy would be a welcome addition to the literature.

Acknowledgments

This research was supported by grant number DMS-2150343 from the National Science Foundation, and was conducted during the Lafayette College Mathematics Department Research Experience for Undergraduates Program during Summer 2023.

References

[1] Turnout in U.S. has soared in recent elections but by some measures still trails that of many other countries, Pew Research Center (November 2022), available online at the URL: `https://www.pewresearch.org/short-reads/2022/11/01/turnout-in-u-s-has-soared-in-recent-elections-but-by-some-measures-still-trails-that-of-many-other-countries/`.

[2] E. Bonabeau, Agent-based modeling: methods and techniques for simulating human systems, *PNAS USA*, **99** (2002), 7280–7287.

[3] S.F. Railsback, V. Grimm, *Agent-based and individual-based modeling: a practical introduction*, Princeton University Press, 2019.

[4] J.G. Gimpel, J.J. Dyck, D.R. Shaw, Registrants, voters, and turnout variability across neighborhoods, *Polit. Behav.*, **26** (2004), 343–375.

[5] I.S. Foladare, The effect of neighborhood on voting behavior, *PSQ*, **83** (1968), 516–529.

[6] W. Minozzi, H. Song, D.M.J. Lazer, M.A. Neblo, K. Ognyanova, The incidental pundit: who talks politics with whom, and why?, *Am. J. Pol Sci*, **64** (2020), 135–151.

[7] L.F. Lafuerza, L. Dyson, B. Edmons, A.J. McKane, Simplification and analysis of a model of social interaction in voting, *Eur. Phys. J. B*, **89** (2016).

[8] , V.K. Singh, S. Basak, N. Modanwal, Agent based modeling of individual voting preferences with social influence, in: D. Nagamalai, E. REnault, M. Dhanuskodi (Eds.), *Trends in Computer Science, Engineering, and Information Technology*, Springer Berlin Heidelberg, 2011.

[9] , I. Moya, M. Chica, J.L. Sáez-Lozano, Ó. Cordón, Simulating the influence of terror management strategies on the voter ideological distance using agent-based modeling, *Telemat. Inform.*, **63**, (2021).

[10] A. Di Benedetto, C.E. Wieners, H.A. Dijkstra, H.T. Stoof, Media preference increases polarization in an agent-based election model, *Physica A Stat. Mech. Appl.*, **626**, (2023).

[11] J.K. Madsen, T.D. Pilditch, A method for evaluating cognitively informed micro-targeted campaign strategies: an agent-based model proof of principle, *PLOS ONE*, **13** (2018), 1–14.

[12] A. Mitra, Agent-based simulation of district-based elections with heterogenous populations, *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, (2023), 2730-2732.

[13] F. Palombi, S. Toti, Voting behavior in proportional elections from agent-based models, *Phys. Procedia*, **62** (2015), 42–47.

[14] L. Qiu, R. Phang, *Agent-based modeling in political decision-making*, Oxford University Press, 2020.

[15] M. Gao, Z. Wang, K. Wang, C. Liu, S. Tang, Forecasting elections with agent-based modeling: two live experiments, *PLOS ONE*, **17** (2022), 1–11.

[16] E. Fieldhouse, L. Lessard-Phillips, B. Edmonds, Cascade or echo chamber? a complex agent-based simulation of voter turnout, *Party Pol.*, **22** (2016), 241–256.

[17] S.S. Gera, Modeling the effects on voter turnout: an agent based approach, Ph.D. Thesis, (2018).

[18] J. Bara, O. Lev, P. Turrini, Predicting voting outcomes in the presence of communities, echo chambers and multiple parties, *Artif. Intell.*, **312** (2022).

[19] F. Justwan, B. Baumgaertner, J.E. Carlisle, A.K. Clark, M. Clark, Social media echo chambers and satisfaction with democracy among democrats and republicans in the aftermath of the 2016 US elections, *J. Elect. Public Opin. Parties*, **28** (2018), 424–442.

[20] M.E. Olsen, Social participation and voting turnout: a multivariate analysis, *Am. Sociol. Rev.*, **37** (1972), 317–333.

[21] H. Evrenk, C.-Y. Sher, Social interactions in voting behavior: distinguishing between strategic voting and the bandwagon effect, *Public Choice*, **162** (2015), 405–423.

[22] A.S. Gerber, T. Rogers, Descriptive social norms and motivation to vote: everybody's voting and so should you, *JOP*, **71** (2009), 178–191.

[23] V.A. Bali, L.J. Robison, R. Winder, What motivates people to vote? the role of selfishness, duty, and social motives when voting, *SAGE Open*, **10** (2020).

[24] J. Harder, J.A. Krosnick, Why do people vote? a psychological analysis of the causes of voter turnout, *J. Soc. Issues*, **64** 2008), 525–549.

[25] A.C. Goldberg, P. Sciarini, Voter turnout in direct democracy: a joint analysis of individual, referendum and community factors, *Eur. J. Polit. Res.*, **62** (2023), 25–46.

[26] D.K. Levine, T.R. Palfrey, The paradox of voter participation? a laboratory study, *APSR*, **101** (2007), 143–158.

[27] E. Borgonovo, M. Pangallo, J. Rivkin, L. Rizzo, N. Siggelkow, Sensitivity analysis of agent-based models: a new protocol, *Comput. Math. Organ. Theory*, **28** (2022), 52–94.

[28] V. Grimm, S. Railsback, C.E. Vincenot, U. Berger, C. Gallagher, D.L. DeAngelis, B. Edmonds, J. Ge, J. Giske, J. Groeneveld, A.S.A. Johnston, A. Milles, J. Nabe-Nielsen, J.G. Polhill, V. Radchuk, M.S. Rohwader, R.A. Stillman, J.C. Thiele, D. Ayllon, The ODD protocol for describing agent-based and other simulation models: a second update to improve clarity, replication, and structural realism, *JASSS*, **23** (2020).

[29] S. Ansolabehere, B.F. Schaffner, CCES Common Content, 2016, (2017).

[30] Pew Research Center, Beyond red vs. blue: the political typology, available online at the URL: `https://www.pewresearch.org/politics/2021/11/09/beyond-red-vs-blue-the-political-typology-2/`.

[31] M. Rolfe, *Voter turnout: a social theory of political participation, political economy of institutions and decisions*, Cambridge University Press, 2012.

[32] U. Wilensky, NetLogo, Center for Connected Learning and Computer-Based Modeling, available online at the URL: `http://ccl.northwestern.edu/netlogo/`.

[33] NetLogo 6.4.0 User Manual, BehaviorSpace guide, available online at the URL: `https://ccl.northwestern.edu/netlogo/docs/behaviorspace.html`.

[34] A. Saltelli, K. Chan, E. Scott, *Sensitivity Analysis*, Wiley, 2009.

[35] R. Smith, *Uncertainty quantification: theory, implementation, and applications*, SIAM Computational Science and Engineering, 2014.

[36] J. Herman, W. Usher, SALib: an open-source Python library for sensitivity analysis, *J. Open Source Softw.*, **2** (2017).

[37] J.H. Kwakkel, M. Jaxa-Rozen, pyNetLogo, available online at the URL: `https://pynetlogo.readthedocs.io/en/latest/`.

[38] U.S. Census Bureau, Table 4a: reported voting and registration, for states, available online at the URL: https://www.census.gov/data/tables/time-series/demo/voting-and-registration/p20-580.html.

[39] Commonwealth of Pennsylvania, Voting and election reports, available online at the URL: https://www.pa.gov/en/agencies/dos/resources/voting-and-elections-resources/election-reports.html#accordion-ea923da40b-item-b01fb1423f.

[40] Department of State, 2016 presidential election statewide official returns, available online at the URL: https://www.electionreturns.pa.gov/_ENR/General/SummaryResults?ElectionID=54&ElectionType=G&IsActive=0.

[41] New York State Board of Elections, Enrollment by county, available online at the URL: https://www.elections.ny.gov/EnrollmentCounty.html.

[42] Ballotpedia, Voter turnout in the 206 pivot counties that voted Obama-Obama-Trump, available online at the URL: https://ballotpedia.org/Voter_turnout_in_the_206_Pivot_Counties_that_voted_Obama-Obama-Trump.

[43] Iowa Secretary of State's Office, Voter registration totals by county, available online at the URL: https://sos.iowa.gov/elections/voterreg/county.html#2016.

[44] Associated Press, Iowa 2016 presidential election results, available online at the URL: https://www.politico.com/2016-election/results/map/president/iowa/.

[45] N. Klein, O. Stavrova, Respondents with more extreme views show moderation of opinions in multi-year surveys in the USA and the Netherlands, *Commun. Psychol.*, **1** (2023).

Grace A. Brophy
Hamilton College
198 College Hill Rd
Clinton, NY 13323
E-mail: gbrophy@hamilton.edu

Audrey R. Rips-Goodwin
University of Kansas
1450 Jayhawk Blvd
Lawrence, KS 66045
E-mail: audrey.rips@ku.edu

Lucy A. Wilson
Bryn Mawr College
101 N Merion Ave
Bryn Mawr, PA 19010
E-mail: lawilson@brynmawr.edu

Allison L. Lewis
Lafayette College
730 High St.
Easton, PA 18042
E-mail: lewisall@lafayette.edu

Received: March 23, 2025 **Accepted**: June 16, 2025
Communicated by Alex Chen