https://pubsonline.informs.org/journal/ijoc

INFORMS JOURNAL ON COMPUTING
Articles in Advance, pp. 1–16
ISSN 1091-9856 (print), ISSN 1526-5528 (online)

The Sensitivity of the U.S. Presidential Election to Coordinated Voter Relocation

Carlos Cardonha,[a] David Bergman,[a,*] Andre A. Cire,[b] Leonardo Lozano,[c] Tallys Yunes[d]

[a] Department of Operations and Information Management, University of Connecticut, Storrs, Connecticut 06828; [b] Department of Management, University of Toronto Scarborough and Rotman School of Management, Toronto, Ontario M1C 1A4, Canada; [c] Operations, Business Analytics & Information Systems, University of Cincinnati, Cincinnati, Ohio 45221; [d] Miami Herbert Business School, University of Miami, Coral Gables, Florida 33146

*Corresponding author

Contact: carlos.cardonha@uconn.edu, https://orcid.org/0000-0002-1439-5205 (CC); david.bergman@uconn.edu, https://orcid.org/0000-0002-5566-5224 (DB); andre.cire@rotman.utoronto.ca, https://orcid.org/0000-0001-5993-4295 (AAC); leolozano@uc.edu, https://orcid.org/0000-0002-9359-5984 (LL); tallys@miami.edu, https://orcid.org/0000-0002-8308-7812 (TY)

Received: October 9, 2024
Revised: March 16, 2025; May 23, 2025; October 16, 2025
Accepted: August 13, 2025
Published Online in Articles in Advance: November 12, 2025

https://doi.org/10.1287/ijoc.2024.0990

Abstract. U.S. presidential elections are determined by the Electoral College. In all but two states, a statewide winner-take-all system for electors can lead to decisive outcomes based on narrow victories in key states. Small groups of voters can significantly impact results, not only through turnout but also through a less-explored mechanism: the strategic relocation of a relatively small number of dedicated voters across state lines. The extent to which election outcomes are sensitive to such coordinated movements has not been thoroughly investigated. We introduce an analytical framework that integrates forecasting, simulation, and optimization to identify these pivotal voter shifts. Our findings show that small-scale relocations can meaningfully alter election probabilities under a range of parameter settings and polling data sources. Furthermore, we examine how the optimization-based recommendations align with actual election results, demonstrating that the suggested movements would have been beneficial in the 2024 U.S. presidential election—even when based on pre-election data. Given the remarkably small number of individuals required and the fact that electoral residency in many states can be established within about a month, our results have direct implications for policymaking and campaign strategy. Moreover, they highlight new opportunities for applying operations research methods to political science.

History: Accepted by Alice E. Smith, Editor-in-Chief.
Funding: This work was supported by an Natural Sciences and Engineering Research Council of Canada Discovery Grant.
Supplemental Material: The software that supports the findings of this study is available within the paper and its Supplemental Information (https://pubsonline.informs.org/doi/suppl/10.1287/ijoc.2024.0990) as well as from the IJOC GitHub software repository (https://github.com/INFORMSJoC/2024.0990). The complete IJOC Software and Data Repository is available at https://informsjoc.github.io/.

Keywords: elections • network models • simulation

1. Introduction

Can a small coordinated interstate movement of voters sway the outcome of a U.S. presidential election? Although this idea is often dismissed due to perceived legal, financial, and logistical challenges, we demonstrate that even a modest, strategically coordinated movement of voters could significantly impact a U.S. presidential election.

Our analysis is rooted in the intricate dynamics of the Electoral College, where in all but two states (Maine and Nebraska) a winner-takes-all allocation of the electoral votes from that state is almost always implemented. The power of the states in determining the outcome of a presidential election is heterogeneous (Rabinowitz and MacDonald 1986), and the current electoral system amplifies the importance of *swing* states. Over the past 25 years, elections have been determined by remarkably narrow differences in the popular vote within these states. One notable example is the 2000 presidential election, won by George W. Bush with 271 electoral votes (versus 266 for Al Gore). In that election, the Republican candidate won Florida's 25 electoral votes by a margin of 537 popular votes—equivalent to a mere 0.005% of the state's eligible voters (Federal Election Commission 2000).

Beyond these thin margins of victory, another key factor to consider is the geographic proximity between swing states and states with more predictable outcomes. For example, the neighboring swing states Nevada and Arizona share a border with California, a Democratic stronghold. Similarly, the swing state of Georgia is bordered almost entirely by Republican strongholds Tennessee, Florida, and Alabama. Referencing the 2000 election again and assuming one had access to perfect information, relocating 538 Democratic voters from Alabama or Georgia to Florida would have changed the outcome of that presidential election. Building on these two factors and through a combination of forecasting, simulation, and optimization, we demonstrate how precisely coordinated, small-scale, short-distance moves could have dramatically influenced the outcome of the 2024 U.S. presidential election.

Voter relocation strategies abound with questions about feasibility in practice. A close inspection of the main potential hurdles indicates that it would be far less challenging to implement than what common perception might suggest. In particular, and perhaps more importantly, current U.S. legislation does not impose barriers on relocation strategies. There is a 30-day maximum limit on residency requirements for voting in federal elections, including presidential elections, based on the 1970 amendments to the Voting Rights Act (42 U.S.C. § 1973aa-1). This federal law stipulates that no state may impose a residency requirement of more than 30 days for voting in any presidential election. Additionally, the U.S. Supreme Court reinforced this in the 1972 case *Dunn v. Blumstein* (405 U.S. 330), where the Court ruled that lengthy residency requirements violated the Equal Protection Clause of the 14th Amendment, despite what some politicians might have called for (Bump and Bronner 2023). Therefore, relocation translates to voting eligibility in a relatively short period.

Another consideration is practicality, as the relocation burden for a voter would be far greater than just voting in their current state of residence. Although turnout initiatives typically focus on motivating the less engaged segments of voters, the result of our study involving relatively few movements suggest that attention can be placed on enthusiastic voters who are already planning to vote; for reference, election rallies frequently have tens of thousands of people in attendance (Peoples et al. 2024). Note that interstate relocations already happen in an organic way; in particular, nearly 8 million people have moved between states in recent years in an uncoordinated fashion (Ismail 2023). Moreover, the 100 companies with the largest workforce in the U.S. employ more than 80,000 people each (Stock Analysis 2024). As a result, even a corporate relocation could inadvertently shift the electoral balance and prove decisive in a closely contested election.

From a financial standpoint, the cost involved with a small number of short-distance relocation movements would be in line with donation to presidential campaigns. For example, as of August 15, 2024, the estimated cost of moving a two-bedroom house from Lowell, MA to Nashua, NH is between $2,081 and $2,542 (Move, Inc 2024). Using the midpoint of this cost range, a rough estimate for the total cost of moving 50,000 people is around $150 million. For reference, more than 200,000 people had donated more than $3,300 to 2024 presidential election campaigns by August 15, 2024 (Open Secrets 2024). Rather than donating to a campaign—where dollars do not necessarily translate to votes (Krasno and Green 2008)—highly motivated voters could instead relocate and directly influence election outcomes. It is important to clarify that the authors neither endorse nor condone these relocation strategies; instead, we aim to shed light on how such planned efforts could impact the democratic process.

Our work contributes to the existing research on elections in the operations research and the political science literature. Specifically, U.S. elections have been studied since the 1960s (Hess et al. 1965, Garfinkel and Nemhauser 1970), with emphasis on local (county-level and state-level) elections. Special attention has been dedicated to political districting and gerrymandering (Validi and Buchanan 2022, Validi et al. 2022, Swamy et al. 2023), examining how electoral boundary design can skew representation and affect the fairness and outcomes of local elections. Fairness considerations also arise in the composition of citizens' assemblies for public policy deliberations, where descriptive representation (matching quotas on attributes such as age, gender, and education) can conflict with democratic equality (each individual should have an equal chance to serve; Flanigan et al. 2021). Another line of research addresses the location and consolidation of polling places to promote voter participation (Haspel and Knotts 2005, Schmidt et al. 2024). For presidential elections, however, state-level popular votes are unaffected by these internal subdivisions (except in Maine and Nebraska). The interest in election forecasts has also grown significantly over the years (Kaplan and Barnett 2003, Hummel and Rothschild 2014), especially after numerous polls failed to predict the outcome of the 2016 election, won by Donald Trump (Wright and Wright 2018). Additional applications of optimization to political problems include the allocation of resources to maximize seats in parliament (Güney 2018), estimating the minimal fraction of the popular vote necessary to elect a president in the Electoral College (Belenky 2008), designing test decks for logic-and-accuracy testing of voting machines (Crimmins et al. 2025), and composing citizens' assemblies under fairness constraints (Flanigan et al. 2021).

2. Problem Description

We investigate the problem from the perspective of the two main parties in the U.S. political system. We refer to them as p_1 and p_2 in our model, and assume without loss of generality that we are solving the problem for party

p_1. Let $\mathcal{U}$ be the set of electoral units, which is a partition $\mathcal{U} := \mathcal{S} \cup \mathcal{D}$ of state units $\mathcal{S}$ and district units $\mathcal{D}$. Each U.S. state has a one-to-one mapping to state units in $\mathcal{S}$. In addition, Maine and Nebraska are associated with two and three district units, respectively, in $\mathcal{D}$. Each unit u is assigned $v_u \in \mathbb{Z}_+$ electoral votes, with $\sum_{u\in\mathcal{U}} v_u = 538$ for the 2024 U.S. presidential election. A candidate who wins the popular vote at a unit u receives all v_u votes.

Each unit $u \in \mathcal{U}$ is composed of counties $\mathcal{C}_u$, where $\mathcal{C} := \cup_{u\in\mathcal{U}} \mathcal{C}_u$ is the set of U.S. counties. Each county belongs to exactly one state unit, and all votes in a county $c \in \mathcal{C}_s$ count toward the state s. A county c in the states of Maine and Nebraska belongs to at least one district unit C_d, $d \in \mathcal{D}$. Let $f_{c,d} \in [0,1]$ be the fraction of county c's population within district unit $d \in \mathcal{D}$. We assume that each vote in a county $c \in \mathcal{C}_d$ counts for $f_{c,d}$ votes toward the district $d \in \mathcal{D}$.

We wish to identify a movement matrix $\mathbf{X} \in \mathbb{Z}_{\geq 0}^{|\mathcal{C}|\times|\mathcal{C}|}$ to most effectively increase the probability of p_1 obtaining at least K electoral votes, where $x_{i,j}$ is the number of people relocating from county c_i to county c_j; we set K to 270 in our study, as this was the minimum number of electoral votes needed to win the 2024 U.S. presidential election. We assume that each movement involves identifiable, highly engaged electors of p_1, that is, individuals who will vote in the elections and choose p_1. Specifically, let $\tilde{N}_{p_1,u}$ and $\tilde{N}_{p_2,u}$ be random variables representing the number of votes parties p_1 and p_2 receive in unit u, respectively. We consider the following stochastic problem:

$$
\begin{aligned}
\max_{\mathbf{X}\in\Omega} \quad & \mathbb{P}\left(\sum_{u\in\mathcal{U}} v_u \tilde{W}_u \geq K\right) \\
\text{s.t.} \quad & \tilde{W}_u = \mathbb{I}(\tilde{T}_u > \tilde{N}_{p_2,u}) && \forall u \in \mathcal{U}(= \mathcal{S} \cup \mathcal{D}) \\
& \tilde{T}_s = \tilde{N}_{p_1,s} + \sum_{c_i\in\mathcal{C}_s} \sum_{c_j\in\mathcal{C}\setminus\mathcal{C}_s} (x_{j,i} - x_{i,j}) && \forall s \in \mathcal{S} \\
& \tilde{T}_d = \tilde{N}_{p_1,d} + \sum_{c_i\in\mathcal{C}_d} f_{c_i,d} \sum_{c_j\in\mathcal{C}\setminus\mathcal{C}_d} (x_{j,i} - x_{i,j}) && \forall d \in \mathcal{D}.
\end{aligned}
\tag{P}
$$

In formulation $\mathbf{P}$, $\tilde{T}_u$ is a random variable that denotes the number of votes received by p_1 in unit u after the relocations in and out of all counties in $\mathcal{C}_u$, as defined in the second and third constraints. Party p_1 obtains all v_u electoral votes if $\tilde{T}_u > \tilde{N}_{p_2,u}$, which here is modeled as a Bernoulli random variable $\tilde{W}_u$ in the first constraint. The objective maximizes the probability that p_1 secures at least K electoral votes. Finally, we use set Ω to indicate that the set of feasible solutions to $\mathbf{P}$ may be subject to additional constraints, such as limits on the number of people moving in and out of a county (Section 5.3).

3. Analytical Framework and Models

We identify voter relocation strategies in a three-step process that considers the stochastic nature of the problem. First, we design a simulation model to generate scenarios of voter turnout by party and county. Next, we calibrate these scenarios using a Bayesian-style update with predictions from polling aggregation services. Finally, we apply an optimization model based on sample average approximation (SAA) (Kleywegt et al. 2001) that processes scenarios from the simulation model to identify county-to-county voter movements. The optimization is independent from the simulation model, so any alternative simulation model generating voter turnout per party and county can substitute the one presented in this work.

3.1. Voter Turnout Simulation Model

The simulation model was designed to capture historical voter turnout per county, consider correlation in voter turnout between counties, and reflect recent voting preference trends observed in polling data. For this, we use growth models derived from county-level voting returns from previous elections. Our model is designed to capture a well-known spatial correlation across counties under partisan stability across the years (Kim et al. 2003, Fiorino et al. 2022, Bump and Bronner 2024). Specifically, the number of voters per county per candidate is modeled as a collection of correlated lognormal distributions that defines the baseline for simulating voter turnout.

For each party p, county c, and election year t, let $v_{p,c,t}$ be the number of people who voted for p in c in the election year t. We define the number of voters for the party p in county c as $\tilde{V}_{p,c} \sim \mathcal{LN}(\mu_{p,c}, \sigma_{p,c})$, that is, a lognormal distribution with the normal component having mean $\mu_{p,c}$ and standard deviation $\sigma_{p,c}$. The lognormal distribution was chosen because (a) it is positive, (b) captures counties with small voting populations, (c) correlates variables based solely on historical trends, and (d) accounts for heteroskedasticity observed in our data.

Finally, we note that the discrete outcomes (number of votes) are sufficiently large so that a continuous distribution is appropriate.

Given an arbitrary ordering of the counties and parties, we let $\boldsymbol{\mu}$ and Σ be the vector of the expected values and the covariance matrix of the normal components of the number of votes per county and party, respectively. We estimate them as $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$ using the following procedure. We model voter turnout using estimates of the average growth rate $\hat{g}_{p,c}$ for each party p and county c based on historical voter data for the previous T elections as follows:

$$\hat{g}_{p,c} := \frac{1}{T}\sum_{t=1}^{T}\frac{(v_{p,c,t+1}) - (v_{p,c,t})}{v_{p,c,t}}.$$

Based on the growth rates above, the expected turnout for the election of interest is estimated as

$$\hat{\mu}_{p,c} := \mathbb{E}[\ln(\tilde{V}_{p,c})] = \ln(v_{p,c,t-1} \cdot \hat{g}_{p,c}).$$

We derive $\hat{\Sigma}$ by computing the sample covariance of the log-transformed voter turnout (for each p and c in the same order as they appear in $\boldsymbol{\mu}$) from previous elections. If the resulting matrix is not positive semidefinite, we replace it with the nearest symmetric positive semidefinite matrix based on the Frobenius norm (Higham 1988). Given $\hat{\boldsymbol{u}}$ and $\hat{\Sigma}$, we generate z scenarios of voter turnout as follows:

1. (Normal Data Simulation) We sample z matrices $\mathbf{A}_q \in \mathbb{R}^{z\times(2|C|)}$ for $q \in [z] := \{1,\ldots,z\}$ from the distribution $\mathcal{N}(\hat{\mathbf{u}}, \hat{\Sigma})$. Each sample $a_{q,p,c}$ represents the log-transformed voter turnout of party p in county c sampled in scenario $q \in [z] := \{1,\ldots,z\}$.
2. (Lognormal Rescale) We exponentiate the adjusted normal data to recover the sampled voter turnout $\xi'_{q,p,c}$ for each party, county, and scenario, that is, $\xi'_{q,p,c} := e^{a_{q,p,c}}$.

The matrix $\boldsymbol{\Xi}' \in \mathbb{R}_+^{z\times 2\times|\mathcal{C}|}$ is the result of the process described above, which includes sampled data describing voting turnout in each scenario. We refer to $\boldsymbol{\Xi}'$ as the *calibration set*.

3.2. Bayesian-Style Update Model

The voter turnout simulation model described in Section 3.1 is based on historical voter trends and does not account for up-to-date information available from recent polls and prediction services. To incorporate this data, we develop a Bayesian-style update to transform the calibration set $\boldsymbol{\Xi}'$ into an *optimization set* $\boldsymbol{\Xi}$ such that the proportion of simulations in which a party wins a state matches a given prediction. This section focus on the case with two parties, as our study is dedicated to the U.S. elections; a generalization of this model encompassing three or more parties is presented in Section 7.2.

Formally, consider a state unit $s \in \mathcal{S}$ for which a party's exogenous winning probability θ_s is greater than the winning probability of that state over all simulations in $\boldsymbol{\Xi}'$ by more than 0.01. We refer to such a party as the *increasing party* $p_\uparrow$, as its number of votes must increase for the winning probability to match θ_s. The adversarial party, in turn, is referred to as the *decreasing party* $p_\downarrow$. Our goal is to switch the smallest fraction λ_s^* of votes in state s from $p_\downarrow$ to $p_\uparrow$ across all simulations in $\boldsymbol{\Xi}'$ to obtain the optimization set $\boldsymbol{\Xi}$. In particular, we wish to minimize λ_s^* to move the fewest number of people to bring the estimates of win probability to match that of the polling site; we refer to λ_s^* as *Bayesian update factor*. We assume a uniform redistribution of voters across each state, meaning the same fraction λ_s^* of voters "flips" in each county of s. We use the following mixed-integer programming formulation to identify λ_s^*:

$$\lambda_s^* := \min_{\lambda,\mathbf{w}} \quad \lambda \tag{1a}$$

$$\text{s.t.}\quad y_q^{p_\uparrow} = \sum_{c\in\mathcal{C}_s}\xi'_{q,p_\uparrow,c} + \lambda\sum_{c\in\mathcal{C}_s}\xi'_{q,p_\downarrow,c} \qquad \forall q \in [z], \tag{1b}$$

$$y_q^{p_\downarrow} = (1-\lambda)\sum_{c\in\mathcal{C}_s}\xi'_{q,p_\downarrow,c} \qquad \forall q \in [z], \tag{1c}$$

$$y_q^{p_\uparrow} + M(1-w_q) \geq y_q^{p_\downarrow} + w_q \qquad \forall q \in [z], \tag{1d}$$

$$\frac{1}{z}\sum_{q\in[z]} w_q \geq \theta_s \tag{1e}$$

$$\lambda \in [0,1], w_q \in \{0,1\}, y_q^{p_\uparrow} \geq 0, y_q^{p_\downarrow} \geq 0, \qquad \forall q \in [z].$$

In addition to variable λ, which is minimized in (1a), the model uses auxiliary variables $y_q^\uparrow$ and $y_q^\downarrow$ to represent the updated number of votes for each party and county in the qth scenario. Moreover, Formulation (1) uses binary variables w_q to indicate whether the increasing party wins s in scenario q. Constraints (1b) and (1c) set the values of $y_q^{\uparrow}$ and $y_q^{\downarrow}$ based on λ. Constraints (1d) set each binary variable $w_q = 1$ if and only if $p_{\uparrow}$ receives more votes than $p_{\downarrow}$ in the qth scenario; M is a sufficiently large constant. Lastly, (1e) asserts that our win probability estimate for $p_{\uparrow}$ matches the prediction θ_s. We remark that this constraint can be tightened by replacing θ_s on the right-hand side with $\lceil z\theta_s \rceil$ because we are dealing with a discrete number of scenarios.

Given an optimal solution λ_s^*, we update the voter turnout scenarios in Ξ' for state s as follows:

$$\xi_{q,p_{\uparrow},c} = \xi'_{q,p_{\uparrow},c} + \lambda_s^* \xi'_{q,p_{\downarrow},c} \quad \forall c \in \mathcal{C}_s,\ \forall q \in [z],$$
$$\xi_{q,p_{\downarrow},c} = (1-\lambda_s^*)\xi'_{q,p_{\downarrow},c} \quad \forall c \in \mathcal{C}_s,\ \forall q \in [z].$$

Section 7.2 discusses a generalization of Formulation (1) for switches involving an arbitrary (three or more) number of parties. In the specific case of the U.S. elections, such a generalized model would allow for switches involving candidates from alternative parties that are aggregated in our model.

3.3. Network Flow Model

Let $L \in \mathbb{N}$ be a fixed upper limit on the number of relocation movements. We identify an optimal relocation strategy based on a given optimization set Ξ using the following network flow model:

$$\max_{\mathbf{x},\mathbf{y},\mathbf{w},\hat{\mathbf{w}}} \ \frac{1}{z}\sum_{q\in[z]} \hat{w}_q \tag{2a}$$

$$\text{s.t.} \quad y_{s,q} = \sum_{c_i \in \mathcal{C}_s}\left(\xi_{q,1,c_i} + \sum_{c_j \in \mathcal{C}\setminus\mathcal{C}_s}(x_{j,i} - x_{i,j})\right) \quad \forall s \in \mathcal{S},\ \forall q \in [z], \tag{2b}$$

$$y_{d,q} = \sum_{c_i \in \mathcal{C}_d} f_{c_i,d}\left(\xi_{q,1,c_i} + \sum_{c_j \in \mathcal{C}\setminus\mathcal{C}_d}(x_{j,i} - x_{i,j})\right) \quad \forall d \in \mathcal{D},\ \forall q \in [z], \tag{2c}$$

$$y_{u,q} + M(1 - w_{u,q}) \ge \sum_{c_i \in \mathcal{C}_u} \xi_{q,2,c_i} + w_{u,q} \quad \forall u \in \mathcal{U},\ \forall q \in [z], \tag{2d}$$

$$\sum_{u\in\mathcal{U}} v_u w_{u,q} \ge K\hat{w}_q \quad \forall q \in [z], \tag{2e}$$

$$\sum_{(i,j)\in\mathcal{A}} x_{i,j} \le L, \tag{2f}$$

$$\sum_{c_j\in\mathcal{C}} x_{i,j} \le \min_{q\in[z]}\{\xi_{q,1,c_i}\} \quad \forall c_i \in \mathcal{C}, \tag{2g}$$

$$\sum_{c_j\in\mathcal{C}} x_{i,j} \le \delta n_i \quad \forall c_i \in \mathcal{C}, \tag{2h}$$

$$\sum_{c_j\in\mathcal{C}} x_{j,i} \le \delta n_i \quad \forall c_i \in \mathcal{C}, \tag{2i}$$

$$y_{u,q} \ge 0,\ w_{u,q} \in \{0,1\} \quad \forall u \in \mathcal{U},\ \forall q \in [z]$$
$$\hat{w}_q \in \{0,1\} \quad \forall q \in [z]$$
$$x_{i,j} \ge 0 \quad \forall c_i, c_j \in \mathcal{C}.$$

Variable $y_{u,q}$ represents the number of votes for party p_1 in unit u under the qth scenario after the relocations. Binary variable $w_{u,q}$ equals one if and only if p_1 wins the popular vote in u under q. Similarly, binary variable $\hat{w}_q$ indicates if p_1 wins the election in q. Each variable $x_{i,j}$ represents the (nonnegative) number of people moved from county c_i to county c_j. Objective (2a) maximizes the estimated probability of p_1 obtaining at least K votes. Constraints (2b) and (2c) count the votes for each state and district, respectively, per scenario after the relocations; recall that $f_{c_i,d}$, used in (2c), is the fraction of county c_i within district unit d. Constraints (2d) and (2e) set the activation variables indicating the victory per state and nationwide, respectively. Constraint (2f) limits the total number of movements to L voters. Constraints (2g) limit the total number of movements out of each county to the minimum number of voters in that county across all scenarios, whereas Constraints (2h) and (2i) limit the total number of movements in and out of each county to a fraction $\delta \in [0,1]$ of the county's population n_i. The values n_i are estimates on the county population for each county using data from the U.S. Census Bureau (2024). These data provide actual county populations from 2020 and estimates of county populations for 2021 and 2022. For each county, we take the average of the growth rates from 2020 to 2021 and from 2021 to 2022 and then multiply the 2022 county populations by that value twice to estimate the county population in 2024.

3.4. Out-of-Sample Evaluation Sets

To ensure that the probability of winning is not overly tied to the optimization set Ξ, we use the same simulation procedure to generate a fresh set Ξ^* of *evaluation* scenarios to assess solution quality, that is, almost surely $\Xi \cap \Xi^* = \emptyset$. Namely, whereas Ξ is used within the network flow model (2), Ξ^* is used exclusively to evaluate the quality of the prescribed movements. With that, we have an out-of-sample prediction for the probability that each candidate will win.

3.5. Practical Considerations and Refinements

We incorporate additional aspects into our model for practicality. First, we ensure that relocation distances are limited to counties located at most 100 miles apart. Moreover, we only consider movements into *swing states*, which we define as those states for which the winner has at most a 70% chance of winning. We also implement a postoptimization routine that tries to reduce the total distance traversed by all movements while maintaining the same aggregate flows out of each state and the same aggregate flows into each state as in the original solution, noting that such modifications do not change the estimated win probability.

With respect to the Bayesian-style update, we use a modified model to account for district units in Maine and Nebraska. Instead of defining a single multiplier for the whole state, we consider two multipliers per district, each denoting the switch of voters between the parties in one district to the other. To ensure consistency, we impose constraints requiring that the updated estimated probabilities for both the state and its districts remain within a small tolerance of the exogenous winning probability. Finally, the objective is to minimize the sum of all the multipliers, again aiming to switch the least amount of people.

4. Simulation Model Validation

We present validation procedures conducted to ensure that the simulations are sufficiently accurate. We first detail our specific implementation, give reasons for the adequacy of the log-normality assumption, and then show how our simulations were quite accurate when compared with actual voter turnout in 2024, on the county level. We complete the section with a discussion on the magnitude of the Bayesian update factors.

4.1. Implementation Details

We trained a voter turnout simulation model (see Section 3.1) that simulates the number of votes received by each candidate in each of the 3,150 counties participating in the 2024 U.S. presidential election using county-level voting returns for the last five elections (2004–2020) with data from the MIT Election Data Science Laboratory (https://electionlab.mit.edu/). All voters for candidates other than the primary Republican or Democratic candidate are grouped into a single "other" category. Thus, our model consists of 9,450 lognormal random variables, where 6,300 variables are associated with the two main candidates.

Recall that our Bayesian-style update model relies on recent poll data. Predictions about election outcomes vary across polling services, so we use three sources in this study:

1. `NS`: "Silver bulletin 2024 presidential election forecast" (Silver and McKown-Dawson 2024), extracted on October 9, 2024;
2. `HILL`: "The Hill's 2024 Election Center" (The Hill 2024), extracted on November 4, 2024; and
3. `RACE`: "Race to the WH" (Race to the WH 2024), extracted on November 5, 2024.

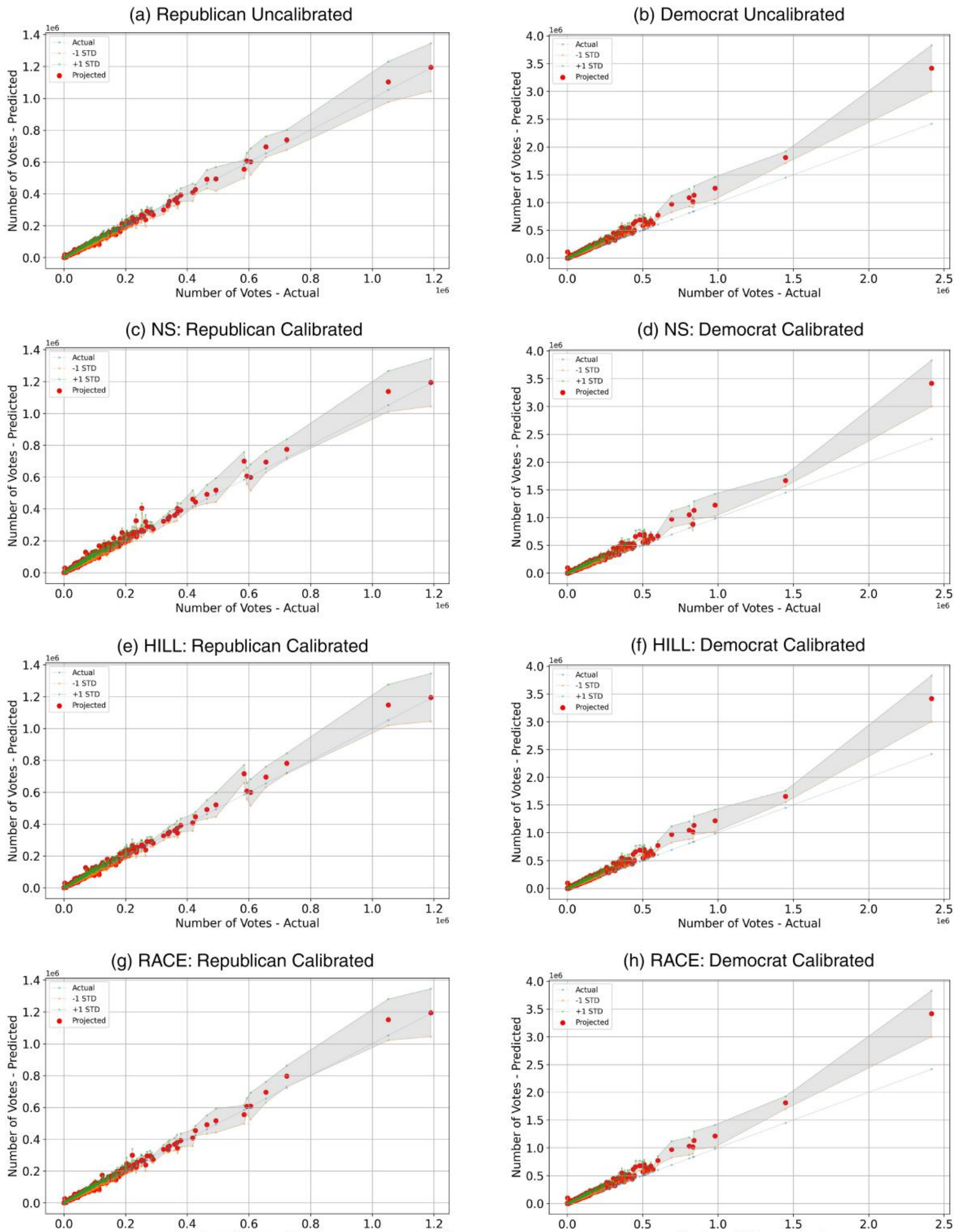
4.2. Adequacy of Log-Normality Assumption

Recall that our model assumes that voter turnout follows a lognormal distribution. We employed a Breusch-Pagan test for heteroskedasticity using 2020 county population data, which yields a Lagrange multiplier of 1,289.73 ($p \leq 0.0001$) when fit to votes for the Democratic candidate. That is, the variance of the residuals is not constant when using the 2020 county population, thereby justifying the use of a distribution that can accommodate such variability.

4.3. Voter Turnout per Party and County

To validate our statistical model, we compare the actual and predicted voter turnouts by county. Figure 1 depicts the actual voter turnout in the 2024 election and a distribution of the predicted voter turnout scenarios, using

Figure 1. (Color online) Comparison of Calibrated and Uncalibrated Models Across Polling Sites



evaluation sets derived from the three polling services. Specifically, we produce two scatter plots with the calibration sets (one per party) plus two scatter plots per polling site, which show the actual number of votes for the Republican and Democratic candidates (x axis) versus the average of the number of voters per county observed in our simulations (y axis). We also include a shaded region in the y axis to show plus-and-minus one standard deviation from the mean over our 5,000 simulations in the evaluation set. Lastly, Table 1 shows the accuracy of uncalibrated and calibrated scenarios across different pool services.

Table 1. Percentage of Scenarios with Turnout Predictions Within One Standard Deviation for Calibrated and Uncalibrated Models, Split by Party and Polling Site

	Republican (R)		Democrat (D)	
Pool service	Calibrated	Uncalibrated	Calibrated	Uncalibrated
NS	78.3%		74.8%	
HILL	79.1%	86.4%	76.3%	70.9%
RACE	81.3%		75.7%	

The results show that the voter simulation model was accurate for the 2024 U.S. presidential election, with most counties having voter turnout within one-standard deviation of the mean. The precalibrated Republican model was the most accurate, with 86.4% of the simulations being within one standard deviation versus only 70.9% for Democrats. In contrast, the accuracy of the Republican model degrades after the Bayesian-style update, whereas the Democrat model becomes better aligned with the actual turnout and closely equates the errors for both models.

This behavior becomes clear when examining Figure 2, which illustrates the progression of voter turnout for the Republican and Democratic candidates in the past six U.S. presidential elections (Leip 1993). Whereas Republican candidates have seen a relatively steady increase in support, Democratic candidates have exhibited more variability, with a notable surge in 2020. The reliance on historical voter turnout led to general overestimation of votes for the Democratic candidate. However, the calibration process ensured that actual voter turnout remained within one standard deviation of the mean for 75%–80% of counties across all prediction pools, aligning with expectations for normal random variables and other common distributions.

4.4. Magnitude of Bayesian Update Factors

Figure 3 shows a histogram of the magnitude of the Bayesian update factors per electoral unit across the three polling services. One can readily see that modest updates are needed to calibrate to live polling data. The average Bayesian update factors are 0.041, 0.025, and 0.013, for NS, HILL, and RACE, respectively. This shows that, although polling updates are necessary, the baselines from the population growth model are not heavily misaligned.

5. Optimization Results

This section presents our optimization results. We describe our implementation details and the elevated win probability resulting from the movements prescribed by our optimization models and evaluate the impact of varying levels of constraints on the change in win probability. All data used in this section were from before the 2024 U.S. Presidential Election. Our data and code are available on GitHub (Cardonha et al. 2025).

Figure 2. (Color online) Popular Vote for Republicans and Democrats in the Last U.S. Presidential Elections

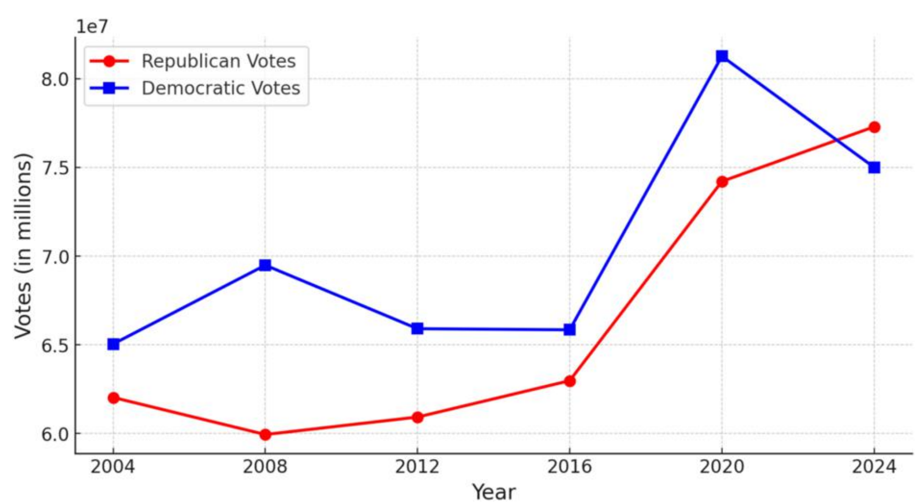
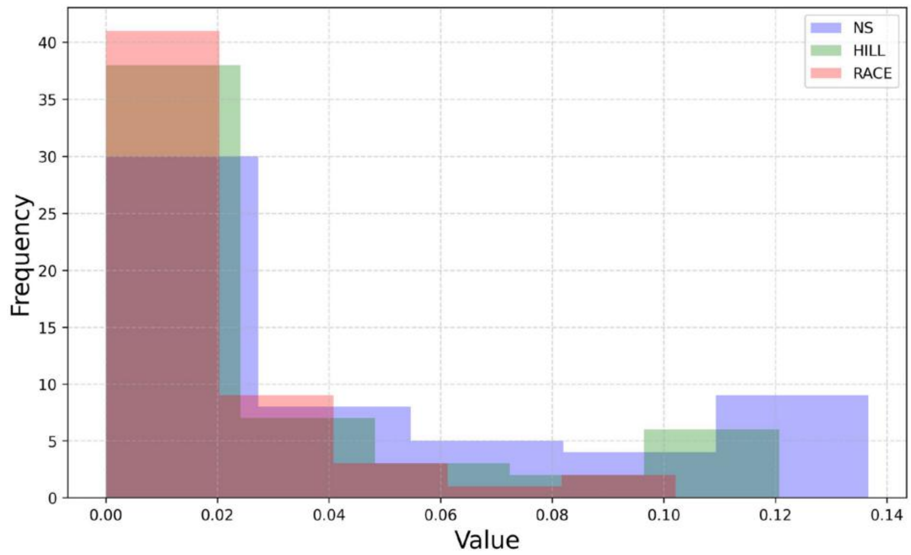


Figure 3. (Color online) Histogram of Bayesian Update Factors Across Polling Sites



5.1. Implementation Details

Our code was implemented in Java. We ran our experiments on an Intel(R) Xeon(R) CPU E5-1650 v4 at 3.60 GHz with 32 GB of RAM. We used `CPLEX 20.1` with default settings to solve all the MIPs.

Movements into and out of a few states were prohibited, either for simplicity or due to data challenges. Specifically, (a) because of distance, we do not allow movements in or out of Alaska and Hawaii; (b) because of district electoral units, we do not allow movements in or out of Maine and Nebraska; and (c) because of discrepancies in naming conventions, movements in and out of Connecticut were also prohibited. In the voting data for Connecticut, the voting records are broken down by Planning Regions as opposed to CT Counties (https://portal.ct.gov/csl/research/ct-towns-counties?language=en_US). A similar discrepancy is also present in Alaska, noting that movements are restricted in the model anyways.

The distance between two counties is calculated as the distance between the centroid of their cities weighted by the population of their component cities. Specifically, for every city, we first extract the population, latitude, longitude, and county from SimpleMaps (https://simplemaps.com/). Next, we compute the coordinates of the county centroids as the weighted average of the latitude and longitude of the cities in each county, weighted by population size.

For our experimentation, we used 1,000 and 5,000 scenarios in the optimization and evaluation sets, respectively. The limit of 1,000 scenarios for the optimization set was dictated by computational bottlenecks and the 5,000 scenarios for evaluation were used to get a refined estimate for win probability. Given that each scenario is a random outcome, the estimate for the confidence interval for the true population proportion p at 95% confidence has width at most 1.39% according to the classical population proportion confidence interval length of $2\sqrt{p(1-p)/n}$, which provided a fairly narrow range.

5.2. Win Probability Elevation

Table 2 presents the win probability elevations derived from our optimization results. Each row indicates the polling site used to generate the optimization set. The columns report the elevation in estimated win probability over the evaluation sets from the three polling sites, resulting from the prescribed movements for 25,000, 50,000, 100,000, 150,000, and 200,000 voter movements. To evaluate the win probability elevation, we fix the movement decisions, identify the fraction of times the candidate wins in the respective evaluation sets, and compare it with the baseline win probability. The first line of the table shows the premovement win probability estimate by polling site. We optimize and evaluate over all combinations to get a robustness check on the estimated win probability change. Lastly, our network flow model does not allow movements changing any county population by more than 5% in these experiments, that is, $\delta = 0.05$.

The results show significant win probability elevations even for relatively small movements. To put these results in perspective, there are roughly 250 million eligible voters in the United States. A movement of 100,000 people involves approximately 0.04% of the voting population. In particular, across all three polling sites, a

Table 2. Elevation in Win Probability

		Republican validation			Democrat validation		
Optimization site	Number of movements	NS	HILL	RACE	NS	HILL	RACE
Baseline	0	41.7%	49.4%	43.2%	58.0%	49.5%	56.2%
NS	25,000	3.4%	0.4%	1.0%	0.9%	1.1%	1.4%
NS	50,000	7.1%	3.2%	4.7%	1.7%	1.9%	2.6%
NS	100,000	8.3%	4.1%	6.2%	2.8%	2.5%	3.2%
NS	150,000	9.1%	5.9%	8.1%	4.2%	4.4%	5.2%
NS	200,000	11.0%	9.8%	12.3%	5.7%	4.2%	4.3%
HILL	25,000	0.7%	1.0%	1.4%	0.4%	1.1%	1.1%
HILL	50,000	1.9%	3.3%	3.9%	1.6%	2.6%	2.9%
HILL	100,000	4.4%	6.0%	7.6%	2.6%	4.7%	5.2%
HILL	150,000	5.2%	7.4%	8.9%	3.9%	6.1%	8.0%
HILL	200,000	6.7%	9.6%	11.7%	4.6%	7.1%	9.8%
RACE	25,000	1.3%	1.0%	1.8%	1.0%	1.2%	1.5%
RACE	50,000	2.1%	3.2%	4.0%	1.3%	2.4%	2.7%
RACE	100,000	5.2%	6.6%	9.0%	2.5%	4.6%	5.7%
RACE	150,000	6.9%	8.8%	11.0%	3.9%	5.9%	8.0%
RACE	200,000	10.0%	9.7%	13.3%	4.5%	6.9%	10.4%

movement of 100,000 voters in favor of the Republican candidate results in a change in win probability of between 4.1% and 9.0%. For the Democratic candidate, that change is between 2.5% and 5.7%. We note that each of the polling sites identified varying *swing states*, and therefore the resulting movements change based on which optimization set is used; one example is the state of New Hampshire, which qualifies as *swing state* in NS but not in HILL and RACE. However, as can be seen in Table 2, the solutions are robust to the various evaluation sets; that is, the movements are impactful even when the polling data used for optimization is different from the one used for evaluation.

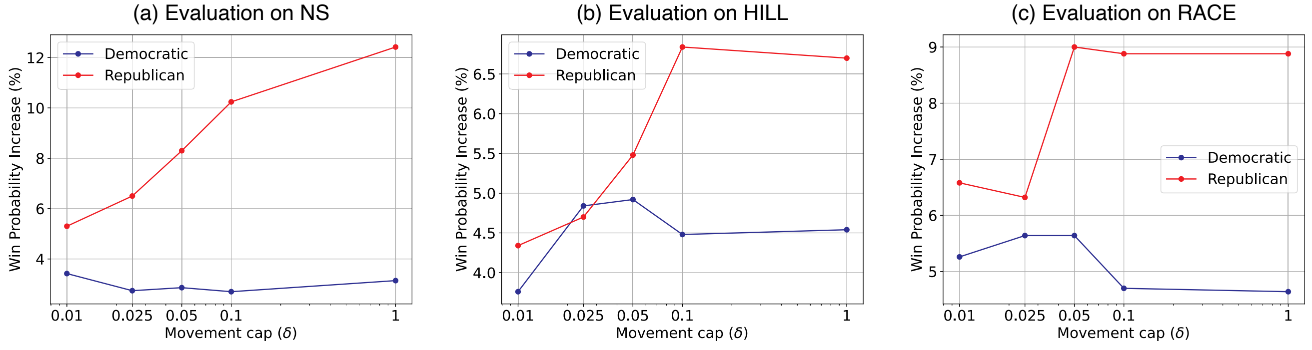
5.3. Sensitivity to Movement Cap

Next, we study the sensitivity of our results to δ, which defines the fraction by which the population of a county may change. Note that arbitrarily large moves would allow for unrealistic or impractical situations. For example, an origin county may be completely depleted from inhabitants, or a destination county may observe a major increase in population size. Additionally, the identification of a large number of enthusiastic supporters in a small county may be challenging. Finally, other logistic limitations, such as housing availability, may hinder large-scale moves.

Figure 4 shows the results of these experiments. We run the modified network flow model for 100,000 movements and parameters $\delta \in \{1, 0.1, 0.05, 0.025, 0.01\}$, using optimization and evaluation sets derived from NS, HILL, and RACE in Figure 4, (a)–(c), respectively. Differences in estimates are reflected in the plots, but the results are overall consistent. In particular, the results show significant gains even under strict limitations on movements. For the Republican candidate, larger values of δ result in higher increases in win probabilities, but we still obtain 5.3% even if $\delta = 0.01$. In contrast, the results for the Democratic candidate are relatively stable, regardless of δ.

A closer inspection of the states used for movements for the Republican and Democratic candidates explain these results. Namely, Republican moves involve states with relatively few counties, so small values of δ impose

Figure 4. (Color online) Sensitivity of Win Probability Increases on δ



significant restrictions. In contrast, Democratic moves involve states with several neighboring counties, so the rearrangement of moves is easier and δ is therefore less impactful. Table 3 shows the largest state-to-state aggregated movements prescribed by the model with no constraints on movements (i.e., $\delta = 1$) for each of the polling sites, for both the Republican and Democratic candidates. We also report the number of counties on the shared border between those states, and the numbers of pairs of counties along their shared borders. This uncovers the discrepancy between the sensitivity of the solutions to δ. For the Republicans, the movements are between states with relatively fewer counties along the shared border, thereby making the constrained movements more impactful.

6. Pre-Election Optimization Results

This section presents the results from our original submission on October 9, 2024, prior to the 2024 U.S. presidential election. Although we have supplemented our analysis in response to the insightful suggestions of the review team and added a collection of results in Section 5 (which are also based only on pre-election data), we retain our original results from the initial submission. We do so for two reasons. First, they offer valuable insights that would have been relevant and informative before the election. Second, they substantiate the findings because there were no election results at the time, thereby solidifying the quality of the models. In these experiments, we solve (2) for each party with the number of relocation movements bounded by 10,000, 15,000, 20,000, 25,000, 50,000, 100,000, 150,000, 200,000, and 250,000 voters. We used `NS` data at the time, using state-by-state predictions as of August 18, 2024. We also adopt an additional postoptimization procedure to avoid excessively small relocation sizes by consolidating county-to-county movements involving fewer than 1,000 people with other movements whenever possible; if the resulting aggregation violates any upper bound constraints, our postoptimization routine simply reverts back to the original solution. Finally, we use $\delta = 1.0$ in these experiments; that is, movements are bounded only by the minimum number of votes in each county across all scenarios.

6.1. Prescribed Moving Strategy

The line plot in Figure 5 shows the increase in win probability that would be realized for each candidate as a function of the number of people moving. Surprisingly, for 10,000 moves, the Republican candidate could already realize a 1.06% increase in election win probability, and with 250,000 moves, an 18.92% increase. For the Democratic candidate, changes are still substantial but less effective: 10,000 and 250,000 moves increase the Democratic win probability by 0.34% and 9.8%, respectively.

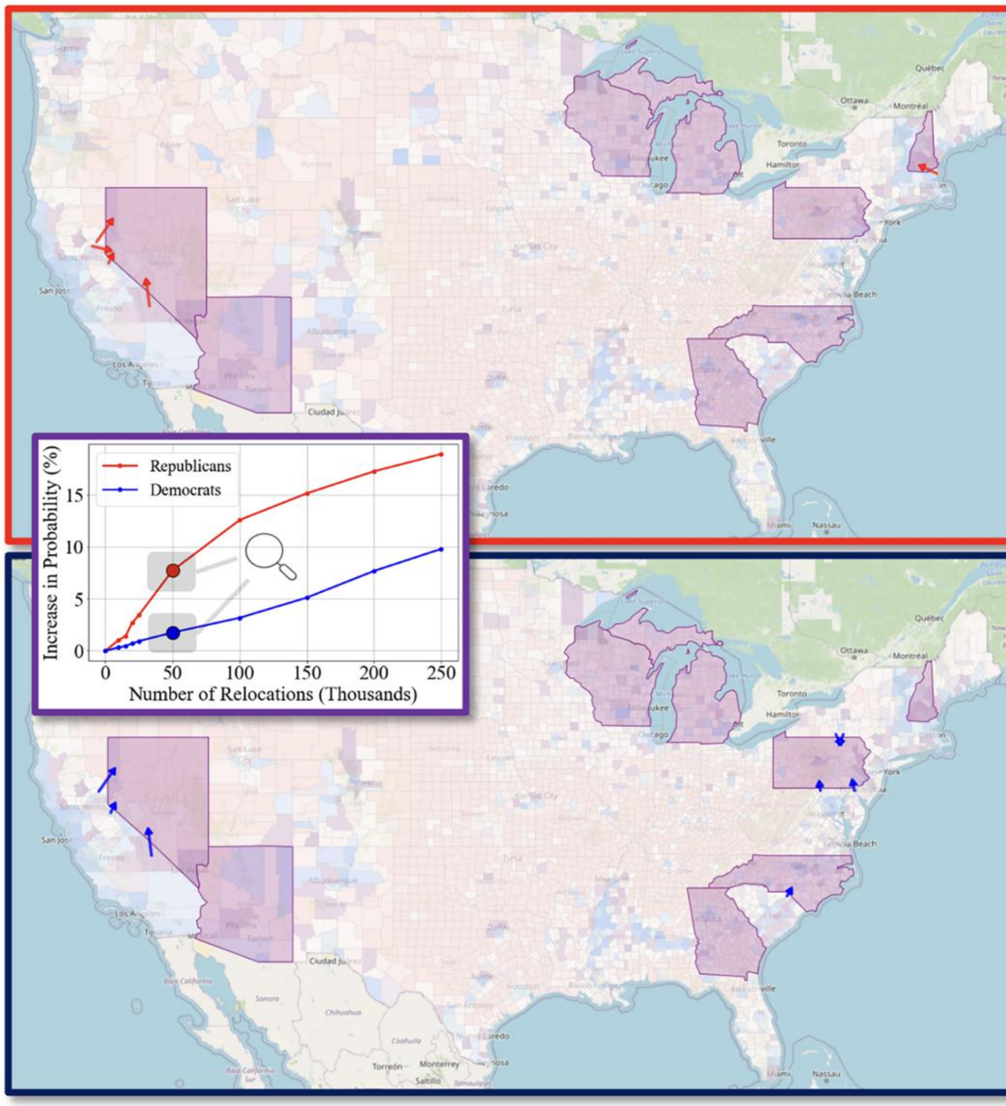
To explain the cause of this disparity between the candidates for the current election, the maps in Figure 5 depict the prescribed relocation strategies for 50,000 people, for both the Republican (top) and the Democratic (bottom) candidates. We observe that 7.72% and 1.76% win probability increases could be attained for the Republican and Democratic candidates in this setting, respectively. For the Republican candidate, the model suggests large movements into New Hampshire (34,755) and Nevada (15,244). For the Democratic candidate, the focus is on Pennsylvania (41,754), accompanied by relatively few moves into Nevada (4,072) and North Carolina (4,173). In this scenario, the average weighted distances traveled by Republican and Democratic supporters are a mere 57.31 and 33.94 miles, respectively.

We hypothesize that the difference in relocation strategy efficiency between the two parties is due to New Hampshire's small size and political leaning based on `NS`. With only four electoral votes and about 800,000 registered voters, a significant shift in New Hampshire's outcome can be achieved with just a few relocations. Specifically, relocating 35,000 Republican voters would increase their candidate's win probability from 24.71% to 73.02%, and considering a total number of movements of 50,000 this still allows additional moves into Nevada, boosting its win probability. In this plan, Republican voters are moving only from Massachusetts and California, states where the Republican candidate is unlikely to win. In contrast, Democrats focus on Pennsylvania, where the candidate has a 55.58% chance of winning. Pennsylvania's larger population of 8.8 million registered voters

Table 3. Largest State-to-State Prescribed Movements with No Constraints (i.e., $\delta = 1$) Along with the Number of Counties on the Shared Border

Candidate	Polling site	Origin	Destination	Movements	Origin counties	Destination counties	Pairs
Republican	`NS`	MA	NH	41,116	4	3	12
Republican	`HILL`	IL	WI	73,307	6	6	36
Republican	`RACE`	IL	WI	93,509	6	6	36
Democratic	`NS`	SC	NC	61,193	11	15	165
Democratic	`HILL`	IL	WI	37,721	6	6	36
Democratic	`RACE`	CA	NV	39,944	9	7	63

Figure 5. (Color online) Movement Patterns for 50,000 People for Republicans (Top) and Democrats (Bottom) and Impact of the Number of People Moved on the Probability of Winning for Different Relocation Sizes (Line Plot)



requires more relocations: 42,000 voters only increase the Democratic candidate's win probability to 61.62%. The remaining 8,000 relocations are insufficient to significantly impact other states, unlike the Republican candidate's dual strategy with New Hampshire and Nevada.

These results are surprising for two reasons. First, they suggest that engaging less than 0.02% of the U.S. voting population is sufficient to increase the Republican candidate's chance of winning by almost 8%. Second, with the same number of movements, the increase in the Democratic candidate's chances of winning do not exceed 2%. Asymmetries like this are not uncommon; for example, rain and snow hurt turnout, and these fluctuations have historically been more beneficial for Republicans (Gomez et al. 2007).

Note that the long-term impact of the movement patterns suggested is complex to estimate. The 50,000-person movement for Republican voters, for example, involves relocating individuals from Democratic strongholds to purple states, potentially achieving a dual effect: flipping a purple state red and increasing the Republican electoral vote count (after a new census) in future elections.

6.2. Impact of the Recommended Solutions

The study identified New Hampshire, Nevada, Pennsylvania, and North Carolina as the most impactful targets at this level of voter relocation. For the Republican candidate, the optimal movements placed voters into New

Hampshire and Nevada. For the Democratic candidate, the analysis recommended directing voters into North Carolina, Nevada, and Pennsylvania. Postelection results highlight the accuracy and significance of our research. Four states—New Hampshire, Wisconsin, Nevada, and Alaska—were ultimately decided by fewer than 50,000 votes (note that we prohibit movements into or out of Alaska). For the Republican candidate, 100% of the suggested movements were into these four states. For the Democratic candidate, only 4,072 of the 50,000 movements were into Nevada, with the majority into Pennsylvania (41,754) and a similar level into North Carolina (4,173).

For the Republican candidate, the focus was on movements into New Hampshire (34,755), where he lost by approximately 22,000 votes. The shared border with Massachusetts makes the movement into New Hampshire perhaps within the scope of reasonability and shows that the model indeed focused on a key state. The prescribed solution also suggested movements into Nevada, where he won, which would have indeed been a good safeguarding measure.

For the Democratic candidate, the largest focus was into Pennsylvania. This state was ultimately lost by approximately 120,000 votes, but it was a critical state to the outcome of the election given its large number of electoral votes (19). No other state with more electoral votes than Pennsylvania's was decided by a smaller margin. We note that it was well understood that Pennsylvania was a critical state, with a Brookings Institute article from October 1, 2024, stating "In all probability, the winner of Pennsylvania will win the election" (Hudak 2024).

7. Extensions

This study examines the U.S. presidential election, which employs a system with several unique characteristics that significantly influenced our algorithmic developments. This section explores the limitations of our modeling assumptions and potential extensions of our algorithms to broader contexts.

7.1. Turnout and Budget of Moved Voters

Model (2) assumes that all relocated voters will participate in the election and vote for the party's candidate. Our analysis focuses on a relatively modest number of voter movements, so identifying a sufficiently large number of enthusiastic supporters—that is, individuals who will reliably vote for a given party—is a plausible assumption for both parties. Nevertheless, we can derive a more conservative variant of (2) parameterized by a turnout parameter $\beta \in [0,1]$, representing the fraction of relocated individuals who actually participate in the election. This modification replaces Equations (2b) and (2c) with the following:

$$y_{s,q} = \sum_{c_i \in \mathcal{C}_s} \left(\xi_{q,1,c_i} + \beta \sum_{c_j \in \mathcal{C} \setminus \mathcal{C}_s} (x_{j,i} - x_{i,j}) \right) \qquad \forall s \in \mathcal{S},\ \forall q \in [z]$$

$$y_{d,q} = \sum_{c_i \in \mathcal{C}_d} f_{c_i,d} \left(\xi_{q,1,c_i} + \beta \sum_{c_j \in \mathcal{C} \setminus \mathcal{C}_d} (x_{j,i} - x_{i,j}) \right) \quad \forall d \in \mathcal{D},\ \forall q \in [z].$$

Similar adjustments apply to other variants where moved people may actually vote for a different party.

One can also extend the network flow model to account for a budget for movements, by replacing (2f) with a knapsack constraint that accounts for moving costs and a unified budget to cover those expenses. Namely, if we denote the cost of moving from county i to county j by $c_{i,j}$ and the unified budget by B, we would replace (2f) with the following:

$$\sum_{(i,j) \in \mathcal{A}} c_{i,j} x_{i,j} \le B.$$

We note that the legal acceptability of the weighted version of the problem with a unified budget is questionable, as a centrally sponsored, coordinated relocation of voters could be interpreted as vote buying.

7.2. Extension to Multiple Parties

The U.S. electoral system is dominated by two political parties. This aspect is important in our analysis because it allowed us to restrict the simulation model to three entities: the two main parties and the aggregation of the others. We discuss possible extensions of our models to handle scenarios where an arbitrary number of parties must be incorporated into the formulations without aggregation.

The extension of our voter turnout simulation model to scenarios with multiple parties is straightforward. As for our Bayesian-style update model, we must consider all possibilities of vote transference. The formulation below extends our model to this case, using the sum of the percentages as objective function.

$$\min_{\boldsymbol{\lambda},\mathbf{w}} \sum_{p_1,p_2} \lambda_{p_1,p_2} \tag{3a}$$

$$\text{s.t. } y_q^p = \sum_{c\in\mathcal{C}_s} \xi'_{q,p,c} + \sum_{p'\in\mathcal{P}\setminus\{p\}} \lambda_{p',p} \sum_{c\in\mathcal{C}_s} \xi'_{q,p',c} - \sum_{p'\in\mathcal{P}\setminus\{p\}} \lambda_{p,p'} \sum_{c\in\mathcal{C}_s} \xi'_{q,p,c} \qquad \forall q\in[z], p\in\mathcal{P}, \tag{3b}$$

$$y_q^p + M(1-w_q^p) \geq y_q^{p'} + w_q \qquad \forall q\in[z], p\in\mathcal{P}, \tag{3c}$$

$$\frac{1}{z}\sum_{q\in[z]} w_q^p \geq \theta_s, \tag{3d}$$

$$\sum_{p\in\mathcal{P}} w_q^p = 1 \tag{3e}$$

$$\lambda_s \in [0,1], w_q \in \{0,1\}, y_q^{p\uparrow} \geq 0, y_q^{p\downarrow} \geq 0, \qquad \forall q\in[z].$$

The constraints and variables in (3) have similar interpretation as in (1). The main differences is that we use variables λ_{p_1,p_2} to indicate the percentage of votes from p_1 moved to p_2. Moreover, binary variables w_q^p indicate the winning party p in simulation q, and equality Constraints (3e) assert that there exists exactly one winner per simulation.

The network flow model focuses on a single party, so its extension to a multiparty scenario requires only modifications in (2d), which must be consider all parties. More precisely, when solving the problem for party p, (2d) is replaced with the following family of inequalities:

$$y_{u,q} + M(1-w_{u,q}) \geq \sum_{c_i\in\mathcal{C}_u} \xi_{q,p',c_i} + w_{u,q} \qquad \forall u\in\mathcal{U},\ \forall q\in[z], p'\in\mathcal{P}\setminus\{p\}.$$

Lastly, multiparty systems would allow interactions between voter movements that could lead to complex outcomes, such as vote-splitting or coalition dynamics, which are absent in the two-party context. There is no trivial extension of the proposed models that considers such aspects, and we believe that the topic is interesting enough to deserve a dedicated study.

7.3. Alternative Electoral Systems

Several features distinguish the U.S. electoral process from other indirect election systems, such as the winner-take-all mechanism. Our study specifically examines the interaction between a population-weighted Electoral College and the winner-take-all assignment method used by nearly all states, a combination that makes election outcomes highly sensitive to small fluctuations in voter turnout within swing states. Extending our results to other electoral systems has technical challenges with potentially smaller impact.

Some electoral systems are inherently complex to model. For instance, the D'Hondt and Sainte-Laguë methods, widely used in Europe, allocate seats using a scaling mechanism that adjusts each party's "bid" for its kth seat. This bid is calculated as the number of votes received by the party divided by a decreasing factor that depends on the number of seats the party has already secured. Specifically, the D'Hondt method divides by $1,2,3,\ldots$, favoring larger parties, whereas the Sainte-Laguë method divides by $1,3,5,\ldots$, offering a relative advantage to smaller parties. In such systems, voter relocation alters these ratios, meaning that solving our problem in this context would require ordering constraints to properly assign seats—a challenge that typically leads to complex mixed-integer linear programming (MILP) formulations. Similar challenges arise in quota-based systems such as the largest remainder method (LRM), used in countries like Brazil, Israel, and Argentina. Under LRM, parties receive seats based on a predefined quota, with any remaining seats allocated in descending order of leftover votes. Modeling the impact of voter relocation in these systems introduces additional constraints, further increasing computational complexity. Additionally, countries like Ireland, Northern Ireland, Scotland, Australia, New Zealand, and India adopt variants of the single transferable vote (STV), in which voters rank candidates. Integrating such systems into a MILP-based framework like ours presents significant difficulties, primarily due to the conditional vote reallocation mechanisms.

In addition to the technical challenges, the impact of voter relocation on alternative electoral systems may not be as striking as those presented in our work. Specifically, the aforementioned electoral systems primarily govern the allocation of seats, and modest voter relocation strategies are unlikely to yield anything more than marginal shifts in expected win probabilities.

8. Conclusions

In this paper, we explore whether a small but strategic movement of people could meaningfully influence the outcome of a U.S. presidential election. Specifically, we approach this question through an analytical lens, leveraging predictive models and operations research methodologies. Although this idea has often been dismissed as impractical, politicians have, in fact, called on people to relocate to vote in key elections (Yang 2020). By building on the structure of the electoral college and examining the proximity of swing states to neighboring states with solid partisan majorities, we demonstrate that even modest population shifts can significantly alter win probabilities.

However, these movements would require careful coordination. One could envision a grassroots effort where motivated voters from critical border counties unite to push their preferred candidate over the finish line. Even more alarming is the prospect of a well-funded malicious actor paying voters to relocate, amplifying their political power in a targeted way. It is important to emphasize that the authors neither endorse nor condone these relocation strategies; instead, we aim to shed light on how such planned efforts could be exploited and their potentially profound impact on the democratic process.

Whether these tactics should be considered a violation of voting laws or merit new legislative measures to detect or prevent such collective actions is beyond the scope of this paper. Nonetheless, it is an issue that warrants further policy discussion and debate. We hope this work will inspire future research in operations modeling and algorithms, contributing to more resilient and equitable election systems.

References

Belenky AS (2008) A 0-1 knapsack model for evaluating the possible Electoral College performance in two-party U.S. presidential elections. *Math. Comput. Modeling* 48(6):665–676.

Bump P, Bronner L (2023) Here's what actually happens when voters move between states. Accessed August 16, 2024, https://www.washingtonpost.com/politics/2023/03/02/marjorie-taylor-greene-national-divorce-partisan-voting/.

Bump P, Bronner L (2024) Where 2024 voters moved since 2020—And how they registered. Accessed August 16, 2024, https://www.washingtonpost.com/politics/2024/04/11/where-2024-voters-moved-since-2020-how-they-registered//.

Cardonha C, Bergman D, Cire AA, Lozano L, Yunes T (2025) The sensitivity of the U.S. presidential election to coordinated voter relocation. https://doi.org/10.1287/ijoc.2024.0990.cd, https://github.com/INFORMSJoC/2024.0990.

Crimmins BL, Halderman JA, Sturt B (2025) Improving the security of United States elections with robust optimization. *Oper. Res.* 73(1):61–85.

Federal Election Comission (2000) 2000 presidential general election results. Accessed August 15, 2024, https://www.fec.gov/resources/cms-content/documents/FederalElections2000_PresidentialGeneralElectionResultsbyState.pdf.

Fiorino N, Pontarollo N, Ricciuti R (2022) Detecting dividing lines in turnout: Spatial dependence and heterogeneity in the 2012 U.S. presidential election. *J. Geovisualization Spatial Anal. (Oxford)* 6(2):34.

Flanigan B, Gölz P, Gupta A, Hennig B, Procaccia AD (2021) Fair algorithms for selecting citizens' assemblies. *Nature* 596(7873):548–552.

Garfinkel RS, Nemhauser GL (1970) Optimal political districting by implicit enumeration techniques. *Management Sci.* 16(8):B495.

Gomez BT, Hansford TG, Krause GA (2007) The Republicans should pray for rain: Weather, turnout, and voting in U.S. presidential elections. *J. Politics* 69(3):649–663.

Güney E (2018) Efficient election campaign optimization using integer programming. *J. Industrial Engrg. Management* 11(2):341–348.

Haspel M, Knotts HG (2005) Location, location, location: Precinct placement and the costs of voting. *J. Politics* 67(2):560–573.

Hess SW, Weaver J, Siegfeldt H, Whelan J, Zitlau P (1965) Nonpartisan political redistricting by computer. *Oper. Res.* 13(6):998–1006.

Higham NJ (1988) Computing a nearest symmetric positive semidefinite matrix. *Linear Algebra Appl.* 103:103–118.

Hudak J (2024) In the presidential election's most important state, the race is a dead heat. Accessed March 10, 2025, https://www.brookings.edu/articles/in-the-presidential-elections-most-important-state-the-race-is-a-dead-heat/.

Hummel P, Rothschild D (2014) Fundamental models for forecasting elections at the state level. *Election Stud.* 35:123–139.

Ismail MS (2023) Number and percentage of state-to-state movers increased between 2021 and 2022. Accessed August 16, 2024, https://www.census.gov/library/stories/2023/11/state-to-state-migration.html.

Kaplan EH, Barnett A (2003) A new approach to estimating the probability of winning the presidency. *Oper. Res.* 51(1):32–40.

Kim J, Elliott E, Wang DM (2003) A spatial analysis of county-level outcomes in U.S. presidential elections: 1988–2000. *Election Stud.* 22(4):741–761.

Kleywegt AJ, Shapiro A, de Mello TH (2001) The sample average approximation method for stochastic discrete optimization. *SIAM J. Optim.* 12(2):479–502.

Krasno JS, Green DP (2008) Do televised presidential ads increase voter turnout? Evidence from a natural experiment. *J. Politics* 70(1):245–261.

Leip D (1993) Atlas of U.S. presidential elections. Accessed March 6, 2025, https://U.S.electionatlas.org/.

Move, Inc (2024) Moving cost calculator for moving estimates. Accessed August 15, 2024, https://www.moving.com/movers/moving-cost-calculator.asp/.

Open Secrets (2024) Donor demographics. Accessed September 15, 2024, https://www.opensecrets.org/elections-overview/donor-demographics.

Peoples S, Catalini M, Kinnard M (2024) Trump rallies 'mega crowd' of approx. 100k at Jersey Shore. Accessed August 15, 2024, https://www.fox5ny.com/news/donald-trump-heading-to-jersey-shore-to-rally-mega-crowd.

Rabinowitz G, MacDonald S (1986) The power of the states in U.S. presidential elections. *Amer. Political Sci. Rev.* 80(1):65–87.

Race to the WH (2024) 2024 election forecasts and polling. Accessed March 4, 2025, https://www.racetothewh.com/.

Schmidt AP, Buell D, Albert LA (2024) Optimal consolidation of polling locations. *Manufacturing Service Oper. Management* 26(3):1028–1042.

Silver N, McKown-Dawson E (2024) Silver bulletin 2024 presidential election forecast. Accessed August 11, 2024, https://www.natesilver.net/p/nate-silver-2024-president-election-polls-model.

Stock Analysis (2024) Biggest U.S. employers. Accessed August 15, 2024, https://stockanalysis.com/list/most-employees/.

Swamy R, King DM, Jacobson SH (2023) Multiobjective optimization for politically fair districting: A scalable multilevel approach. *Oper. Res.* 71(2):536–562.

The Hill (2024) 2024 Election Center. Accessed March 4, 2025, https://elections2024.thehill.com/.

U.S. Census Bureau (2024) County population totals: 2020-2023. Accessed March 12, 2025, https://www.census.gov/data/tables/time-series/demo/popest/2020s-counties-detail.html.

Validi H, Buchanan A (2022) Political districting to minimize cut edges. *Math. Programming Comput.* 14(4):623–672.

Validi H, Buchanan A, Lykhovyd E (2022) Imposing contiguity constraints in political districting models. *Oper. Res.* 70(2):867–892.

Wright FA, Wright AA (2018) How surprising was Trump's victory? Evaluations of the 2016 U.S. presidential election and a new poll aggregation model. *Election Stud.* 54:81–89.

Yang A (2020) An interesting tweet about voter movement. https://x.com/AndrewYang/status/1324809476940189698?