The drop box location problem

Adam Schmidt\textsuperscript{a} and Laura A. Albert\textsuperscript{b}

\textsuperscript{a}Johns Hopkins University Applied Physics Laboratory, Laurel, MD, USA; \textsuperscript{b}Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI, USA

\section*{ABSTRACT}
For decades, voting-by-mail and the use of ballot drop boxes has substantially grown within the USA, and in response, many USA election officials have added drop boxes to their voting infrastructure. However, existing guidance for locating drop boxes is limited. In this article, we introduce an integer programming model, the Drop Box Location Problem (DBLP), to locate drop boxes. The DBLP considers criteria of cost, voter access, and risk. The cost of the drop box system is determined by the fixed cost of adding drop boxes and the operational cost of a collection tour by a bipartisan team who regularly collects ballots from selected locations. The DBLP utilizes covering sets to ensure each voter is in close proximity to a drop box and incorporates a novel measure of access to measure the ability to use multiple voting pathways to vote. The DBLP is shown to be NP-hard, and we introduce a heuristic to generate a large number of feasible solutions for policy makers to select from a posteriori. Using a real-world case study of Milwaukee, WI, U.S., we study the benefits of the DBLP. The results demonstrate that the proposed optimization model identifies drop box locations that perform well across multiple criteria. The results also demonstrate that the trade-off between cost, access, and risk is non-trivial, which supports the use of the proposed optimization-based approach to select drop box locations.

\section*{1. Introduction}
During the 2020 General election within the USA, a record 46\% of voters cast a ballot by mail or absentee in-person (MIT Election Data + Science Lab, 2021). Approximately 41\% of these voters cast a ballot using a drop box (Pew Research Center, 2020\textsuperscript{a}), which are temporary or permanent fixtures similar to United States Postal Service (USPS) postboxes. Many states increased the number of drop boxes during 2020 in response to increased use of the vote-by-mail system and to help mitigate health risks associated with in-person voting (Corasaniti et al., 2020). However, the increase in drop box use is likely not a one time event. The use of non-traditional voting methods within the USA has steadily grown since 1996 (Scherer, 2021). A recent survey of Wisconsin’s election clerks found that approximately 78\% of election clerks would like some use of ballot drop boxes in future elections, and this percentage is higher among clerks from jurisdictions with a large voting age population (Burden, 2021). Many states have since introduced legislation to expand the number of drop boxes available to voters\textsuperscript{1} (Vasilogambros, 2020).

Reasons for casting a ballot using a drop box include the perceived security they offer, anticipated mail delays, and a lack of voter confidence in the USPS (Pew Research Center, 2020\textsuperscript{b}). For many individuals, drop boxes are also in close proximity of their home, work, or daily routine (Stewart III, 2017). Arguably, the primary benefit of drop boxes is the increased accessibility they offer to the voting infrastructure compared with in-person voting. Studies suggest that adding drop boxes to a voting system can increase voter turnout (Collingwood et al., 2018). McGuire et al. (2020) found that a decrease in a mile to the nearest drop box increases the probability of voting by 0.64\%. This finding aligns with the hypothesis of election participation first offered by Downs (1957). According to this hypothesis, potential voters decide whether to vote by comparing the cost (e.g., time) of voting and the potential benefits from voting. It was later argued that voting cost is the significant driver of voter turnout (Sigelman and Berry, 1982; Haspel and Knotts, 2005). We posit that the election infrastructure plays a large role in determining the cost to vote (Collingwood et al., 2018; Cantoni, 2020; McGuire et al., 2020). Thus, if we can improve the accessibility of ballot drop boxes to voters by appropriately designing the drop box infrastructure, then we can increase voter participation, particularly among groups who previously had a high cost to vote and low turnout.

Although drop boxes can increase voter participation, there are many challenges associated with identifying drop box locations and managing the drop box voting system. First, drop boxes can pose a large financial cost. Drop boxes can cost $6000 (Joint COVID Working Group, 2020), and designated video surveillance cameras that increase drop box

\textsuperscript{1}There are challenges to some proposals and even calls to restrict the use of these resources (Vasilogambros, 2020).
security can cost up to $4000 (Schaefer and Gammans, 2020). Second, with an increased number of drop boxes, substantial time and resources must be devoted to collecting ballots. During the election period, it is recommended that bipartisan teams regularly collect ballots (Joint COVID Working Group, 2020). If drop boxes are not strategically placed or if there are a large number of drop boxes, this route may be costly and leave less time to devote to other election tasks. Third, there are security risks associated with ballot drop boxes that must be addressed, although drop boxes are considered reliable (Scala et al., 2022). If the drop box-specific security risks are mitigated appropriately, adding drop boxes to a voting system makes an adversarial attack on the electoral process more challenging. This improves the overall security of the voting system, since the system becomes more distributed (Scala et al., 2022). In addition to the previously mentioned challenges, elections are administered by state and local governments within the USA, and each may have different voting processes. Although the vote-by-mail process is typically similar across different jurisdictions within the USA each jurisdiction may have unique challenges or preferences that necessitates a detailed analysis of potential drop box system design.

In light of these complexities, existing guidelines for selecting drop box locations are often insufficient to support election administrators. In 2020, the Cybersecurity and Infrastructure Security Agency (Joint COVID Working Group, 2020) recommended that a drop box be placed at the primary municipal building, there be a drop box for every 15,000–20,000 registered voters, and more drop boxes should be added where there may be communities with historically low absentee ballot return rates. Some states, such as Michigan, also require locating drop boxes “in an equitable way” (Michigan state, 2022). However, these guidelines are not prescriptive enough to support administrators in identifying an appropriate portfolio of drop box locations that satisfies all the requirements. To our knowledge, the only analytical approach to selecting drop box locations uses a Geographic Information System to determine the locations that served the most voters, allowing for a maximum drive time of 10 minutes (Greene and Ueyama, 2015). This approach overlooks many of the trade-offs within the voting system and ignores socioeconomic differences between voters that may make voting more challenging for some individuals.

Without adequate decision support tools, election administrators may ultimately select drop box locations that perform poorly across multiple criteria by which voting systems are measured. In this article, we propose an Integer Program (IP) to support election administrators in determining how ballot drop boxes should be used in their voting systems when allowed by law\(^2\). We formalize the IP as the Drop Box Location Problem (DBLP). To our knowledge, the DBLP is the first mathematical model of the ballot drop box system to support election planning. The DBLP seeks to minimize the capital and operational cost of the drop box system, ensure equity of access to the voting system, and mitigate risks associated with the drop box system. Loosely, we let access refer to the proximity of the voting infrastructure (e.g., polling places, drop boxes) to voters and the ease with which voters can cast a ballot. Expanding access through the use of drop boxes is an important aspect of the DBLP, since voter turnout is highly correlated with the distance needed to travel to cast a ballot (Cantoni, 2020). We measure access to the drop box voting system using conventional covering sets. In addition, we propose a function based on concepts from discrete choice theory to measure the level of access a voter has to the multiple voting pathways offered by the voting system, including in-person and voting-by-mail pathways, which is a novel aspect of the DBLP.

The remainder of this article is structured as follows. In Section 2, we review the management science literature related to elections. In Section 3, we discuss measures by which the ballot drop box system can be assessed. We then formalize the DBLP and introduce an IP formulation of the DBLP. In Section 4, we discuss solution methods for the DBLP. In Section 5, we introduce a case study of Milwaukee, WI, USA using real-world data. Using this case study, we demonstrate the value of our integer programming approach compared with rules-of-thumb that may otherwise be used. We find that the DBLP outperforms the rules-of-thumb with respect to nearly all criteria considered. We then investigate the trade-off between cost, access, and risk within potential drop box system designs to demonstrate that the trade-off is non-trivial, with the optimization-based approach providing value over ad-hoc approaches. We conclude with a brief discussion in Section 6.

2. Literature review

Much of the management science literature aimed at supporting election planning focuses primarily on in-person voting processes. Some research focuses on identifying and describing the in-person voting process including quantifying the arrival rate of in-person voters, the attrition rate of polling place queues, the check-in service rate, the time to vote, and poll worker characteristics (Spencer and Markovits, 2010). Queueing theory has been widely used to analyze lines at polling locations and identify mitigating practices to avoid long lines (Stewart III and Ansolabehere, 2015; Schmidt and Albert, 2021). Since voting machines have been recognized as a bottleneck in the in-person voting process (Yang et al., 2009), a stream of papers focuses on the allocation of voting machines to polling locations (Allen and Bernshteyn, 2006; Edelstein and Edelstein, 2010; Wang et al., 2015).

Other research focuses on risks of voting systems rather than operational design. The Election Assistance Commission (EAC) (EAC Election and Advisory Board, 2009) analyzes threats to voting processes in the USA for seven voting technology types. Scala et al. (2022) identify security threats for mail-in voting processes and offer a relative score for each to identify the most important threats to address. They identify three drop-box-related threats. First, a misallocation of drop

---

\(^2\)The ability to use or not use drop boxes and in what capacity is typically set by state law.
boxes can suppress voter turnout. Second, a drop box can be damaged or destroyed. Third, ballots within a drop box can be stolen or manipulated. They find the likelihood of drop box risks to be relatively low compared with other risks. Fitzsimmons and Lev (2020) study geographic-based risks by introducing a control problem to study how voter turnout can be manipulated through the strategic selection of polling locations. A few papers attempt to detect disruptions or security incidents following an election (Allen and Bernshteyn, 2006; Highton, 2006).

There are no known papers intended to support election administrators in planning and managing the vote-by-mail system. Our proposed IP addresses the risks of the drop box system (Scala et al., 2022) and employs concepts from the facility location literature. Facility location problems are defined by a set of demands (e.g., voters) and a set of facilities (e.g., drop boxes) that can serve the needs of the demands. Arguably the most widely used facility location model is the Maximal Covering Location Problem (MCLP) (Church and ReVelle, 1974). In the MCLP, a demand is “covered” by, or can be served by, a predetermined set of locations called the covering set. Facility locations are selected to maximize the number of demands covered by at least one facility. The Location Set Covering Problem (LSCP) instead requires that all demands are covered and the cost of the selected facility locations is minimized (Toregas et al., 1971).

The IP introduced within this article extends the Covering Tour Problem (CTP) (Gendreau et al., 1997), which is a variant of the LSCP and the Traveling Salesman Problem (TSP) (Davendra, 2010), by considering additional constraints and objective function terms. These changes allow us to accurately model the drop box voting system. A CTP instance is defined by an undirected weighted graph with two mutually exclusive and exhaustive set of nodes, the tour nodes and coverage nodes. The objective of the CTP is to find a Hamiltonian tour of minimal length over a subset of the tour nodes such that each coverage node is covered by at least one node visited by the tour. The CTP is NP-hard since the TSP can be reduced to it (Gendreau et al., 1997). Several solution methods, including exact (Gendreau et al., 1997; Baldacci et al., 2005) and heuristic (Vargas et al., 2017; Murakami, 2018), have been proposed for the CTP. This article represents the first known application of a CTP variation to voting systems.

### 3. Problem definition

In most states, there are multiple pathways by which voters can cast a ballot, and the accessibility of each pathway can influence voter turnout. Figure 1 illustrates the two main pathways, which are typically divided into “in-person” or “absentee”. With in-person voting, a voter obtains and casts a ballot at their assigned polling location, typically on an election day. With absentee voting, a voter requests a ballot be sent to them and the completed ballot is then returned either through the mail or using a drop box. In some states, voters must provide a reason to vote absentee, whereas in 34 states there is “no-excuse” absentee voting (National Conference of State Legislatures, 2022).

In this article, we are concerned with a sub-pathway of the vote-by-mail process where the voter submits a ballot using a drop box. In this pathway, a voter first requests and receives a ballot through the mail. They then decide to submit a ballot using a drop box rather than through the mail (or not returning it at all). A team of poll workers then collects ballots from the drop boxes, and the ballots are processed at an official election building.

If election administrators decide to add drop boxes, they must decide how many drop boxes to add and where they should be located. The DBLP introduced in this section identifies the optimal placement of drop boxes. The DBLP captures the steps in Figure 1 outlined in red ([:]), since they are the steps that are unique to the drop box system and are influenced by the locations of the drop boxes. Election administrators can use the DBLP during the election planning process to assess the cost, access, and risk of a potential drop box system. This can inform their decision of whether or not to add any drop boxes to the voting system.

### 3.1. Assessing drop box infrastructure

Several performance metrics are used to evaluate voting systems in the USA. The most widely reported election performance metrics are the number of individuals registered to vote and the fraction of eligible voters that cast a ballot, known as voter turnout (MIT Election Data & Science Lab, 2022). Two additional metrics are typically used to assess

---

3Some states, such as Washington, use the “absentee” voting process as their primary voting method. Thus, we use “absentee” loosely in this article, and sometimes refer to it as the vote-by-mail process.
the vote-by-mail system: the proportion of requested ballots that are returned and the number of ballots rejected (MIT Election Data & Science Lab, 2022). The use of ballot drop boxes can lower the rejection rate of mail ballots by reducing the time it takes a ballot to return to election officials. As a result, a voter can be notified of an incorrectly marked ballot more quickly to allow the voter to resubmit their ballot before the election deadline. This is a benefit that we do not explicitly consider in our model. We also posit based on empirical research that a well-designed drop box system can lead to a higher proportion of returned mail ballots and a higher voter turnout by improving the accessibility of the voting infrastructure (Downs, 1957; McGuire et al., 2020).

We elaborate on how access to the voting system is measured. We employ the concept of coverage to measure the access voters have to the drop box system. Under the concept of coverage, a voter covered by a selected drop box location is assumed to have access to the drop box voting system. The locations that provide a voter coverage are called its covering set. Covering sets are flexible and can be defined to account for different modes of transportation, vehicle ownership, and other socioeconomic factors. However, drop boxes are a subcomponent of a larger voting system, and coverage overlooks the access provided by non-drop box voting pathways. In reality, some individuals may have better access to in-person voting than others, and adding drop boxes near them may not substantially benefit them. This necessitates a second measure of access that distinguishes access to the complete voting infrastructure from coverage by the drop box system.

We introduce an access function based on the multinomial/conditional logit model from discrete choice theory (Aloulou, 2018) to capture this dynamic. The application of discrete choice theory to questions within political science is most commonly used to explain or predict choices within a multi-candidate (or party) election (Glasgow and Alvarez, 2008). Discrete choice models have also been used to predict how individuals interact with infrastructure in different application domains. One of the earliest cases of this was the application of a conditional logit model to predict the use of the Bay Area Rapid Transit prior to its construction (Train, 2009). To the best of our knowledge, the current article represents the first use of a function based on discrete choice theory to model access within an optimization model.

The function we introduce makes use of some parameters. Let \( v^0_w > 0 \) be a measure of accessibility in the non-drop box voting system (e.g., in-person polling locations) for voters \( w \). This can be determined, for example, by the distance to the nearest polling location. Let \( a^0_w > 0 \) be a measure of the access that a drop box at location \( n \) would provide to \( w \). This can be determined in part by the proximity of the location to the voters’ places of residence and work and by the various transportation modes available between the voters and the drop box location. Based on empirical studies, the value of \( a^0_w \) should be increasing with decreasing distance (McGuire et al., 2020). Finally, let \( v^1_w > 0 \) be the propensity of \( w \) not to vote. This could be informed by the historical non-voting rate (complement of turnout) or using surveys. Using these parameters, we introduce the following access function to measure the access a group of individuals \( w \) has to all voting pathways, including in-person voting pathways, where \( N^* \) represents the set of selected drop box locations:

\[
A_w(N^*) := \frac{v^1_w + \sum_{j \in N} a^0_{wj}}{v^1_w + v^0_w + \sum_{j \in N} a^0_{wj}}
\]

The access function takes values between zero and one. A value closer to one means that the voting system, including the new ballot drop boxes, is more accessible to individuals \( w \) whereas a value closer to zero means that the voting system is relatively inaccessible to individuals \( w \). In this way, a higher access function value suggests higher turnout for \( w \).

The access function can still be used when a strict interpretation is not reasonable or is not feasible, due to data availability, since the benefit of the access function is a result of its structure. First, the access function models access as a non-binary measure. Second, adding any drop box to the voting system increases the value of the access function but to varying degrees based on the locations of the voter and the drop box. Third, each voter has some heterogeneous level of access to non-drop box voting methods captured by \( v^1_w \), and this access is treated as a constant within the scope of the decision to locate drop boxes. Each voter also has a heterogeneous access function value when no drop boxes are added to the voting system,

\[
A_w(\emptyset) = \frac{v^1_w}{v^1_w + v^0_w},
\]

which is reflective of heterogeneous turnout rates. Fourth, the benefit of adding a drop box near a voter is marginally decreasing as the access function value increases. This incentivises placing drop boxes near populations with low levels of access to other voting pathways.

Although it is desirable to increase voter turnout and access to the voting system, expanding the use of ballot drop boxes may increase the financial cost of managing the election. The costs of the ballot drop box system can be broken into two major groups: fixed or operational. Fixed costs represent the “per drop box” costs such as the initial purchase and costs of securing and maintaining the drop box. Each location may have a different fixed cost, due to varying installation and security equipment requirements. Once drop boxes are installed, jurisdictions incur an operational cost for a bipartisan team to collect ballots from the drop boxes (Joint COVID Working Group, 2020). The operational cost is determined, in part, by the distance between drop boxes, the opportunity cost of the bipartisan team’s time, and the frequency at which the ballots are collected during an election. Since jurisdictions typically use a single
collection team, we assume that a single bipartisan team collects ballots from all drop boxes whenever a collection is conducted, and the drop boxes are visited in an order that minimizes the operational cost, referred to as the collection tour. In addition to introducing new financial costs, drop boxes introduce three types of risks to the voting process that can be mitigated through design requirements. The first risk is that ballot drop boxes can be misallocated in a way that causes voter suppression (Scala et al., 2022). There are two components to this risk. The first is the potential to misallocate drop boxes such that access to the drop box voting system is inequitable. The second is the potential to misallocate drop boxes such that the access to the entire voting system, defined by the multiple voting pathways, is inequitable. These risks are reflected by the number of voters covered by a drop box, using the same definition of coverage introduced earlier, and the value of the access function for each voter, respectively. We can mitigate the risk of voter suppression by requiring that each voter is covered by at least one drop box and that the value of the access function meets some minimal threshold for all voters.

The second risk is that a drop box could be damaged or destroyed (Scala et al., 2022). A nefarious actor could influence an election by targeting drop boxes that provide access to certain voters. The impact of this risk can be mitigated by requiring all voters to be covered by multiple drop boxes, so that voters have redundant access to the drop box system.

The last risk is that ballots submitted to a drop box could be stolen or manipulated. The impact of this risk can be mitigated by ensuring that the collection tour has a low cost. When the collection tour has a low cost, election officials can collect ballots often, leaving fewer at risk. Other implicit design choices also mitigate this third risk. For example, requiring a bipartisan team to collect ballots, rather than one individual, reduces the risk of an insider attack. Likewise, incorporating security-related costs, such as the cost of a video surveillance system, into the fixed cost of a drop box mitigates the risks associated with it.

There are additional risks and mitigations associated with the voting process that are not unique to the drop box infrastructure. For example, there is a risk of an insider attack on ballots stored at an election building after being collected from the drop boxes (Scala et al., 2022). However, these additional risks are outside the scope of the system considered in this article (see Figure 1).

### 3.2. The DBLP

We now formally introduce an IP formulation of the DBLP using the sets, parameters, and variables presented in Table 1.

The DBLP selects drop box locations from a set of potential locations, $N$. Potential drop box locations can be identified using existing guidelines (Joint COVID Working Group, 2020; McGuire et al., 2020). Let $y_n$ be a decision variable that equals one if a drop box is located at location $n \in N$ and zero otherwise. Once drop box locations are selected, a collection tour over them must be found to determine the operational cost of the drop box system. Let $x_{ij}$ be a decision variable that equals one if the collection tour travels between drop box $i$ and drop box $j$, $(i,j) \in E$, and zero otherwise, where $E$ defines the set of undirected edges between each pair of nodes in $N$. That is, for each pair of nodes $i \in N, j \in N$ such that $i \neq j$, $E$ contains either $(i,j)$ or $(j,i)$ but not both. We assume the collection tour always begins and ends at a drop box $s$ located at $s$ (e.g., primary municipal building). Let $T$ represent the locations at which there must be a drop box within our solution (e.g., existing drop box locations). The set $T$ is always non-empty, since $T = \{s\}$ in the extreme case. For each location $j \in N$, let $f_j$ equal the fixed cost of a drop box at $j$. Let $c_{ij}$ represent the operational cost of traveling between drop boxes $(i,j) \in E$ on the collection tour.

Using this notation, we formalize the three goals of the DBLP. The first goal is to minimize the total cost associated with the selected drop box locations. The total cost of the

---

| Table 1. Notation. |
|-------------------|
| **Sets** | |
| $W$ | the set of voter populations |
| $N$ | the potential drop box locations |
| $T \subseteq N$ | the locations at which a drop box must be placed |
| $E$ | all pairs $i \in N, j \in N$ such that $i \neq j$ and $(i,j) \notin E$ |
| $N_w \subseteq N$ | drop box location that cover $w \in W, |N_w| \geq 2$ |
| **Parameters** | |
| $s$ | the start and end of the collection tour |
| $f_j$ | the fixed cost of placing a drop box at location $j \in N$ |
| $c_{ij}$ | the operational cost of traveling between $i \in N$ and $j \in N$ in the collection tour |
| $v_w$ | the propensity of $w \in W$ not to vote |
| $v_w^*$ | the accessibility of the non-drop box voting system to $w \in W$ |
| $g_{ij}$ | the accessibility of location $j \in N$ to $w \in W$ |
| $r$ | minimal allowable value for the access function |
| $q$ | minimal number of drop boxes covering each $w \in W$ |
| **Decision Variables** | |
| $x_{ij}$ | 1 if the collection tour moves between $i$ and $j (i,j) \in E$ and 0 otherwise |
| $y_n$ | 1 if a drop box is placed at location $j \in N$ and 0 otherwise |
The drop box system is the sum of the fixed costs and the cost of the collection tour, $z_1 := \sum_{i,j \in E} c_{ij} x_{ij} + \sum_{j \in N} f_j y_j$. The value of $z_1$ serves as the objective function in the IP formulation of the DBLP.

The second goal of the DBLP is to equitably improve the accessibility of the voting system. Let $W$ denote the collection of voter populations. Let $N_w \subseteq N$ represent the drop boxes that cover $w \in W$. We ensure equitable access to the drop box system by requiring that each voter is covered by $q$ drop boxes. Reasonable values of $q$ are 0, 1, or 2. The cardinality of each covering set must be at least $q$, $|N_w| \geq q$ for all $w \in W$, otherwise the problem is infeasible. We ensure equitable access to all voting pathways by requiring that the access function value is at least $r$ for each $w \in W$, $\min_{w \in W} A_w (N^*) \geq r$ where $N^* = \{ n \in N : y_n = 1 \}$ are the selected drop box locations. This constraint can be viewed as a second objective for the DBLP using the epsilon-constraint approach for multi-objective optimization problems (Mavrotas, 2009).

The third goal of the DBLP is to mitigate the risks associated with the drop box voting system. The risk of misallocating drop boxes in a way that leads to voter suppression is addressed by the second goal of the DBLP. The risk of ballots being susceptible to manipulation once submitted to a drop box is addressed by minimizing the cost of the collection tour, which is captured within $z_2$. The risk of damage to or destruction of drop boxes is a way that degrades voter access to the voting system is mitigated by ensuring each voter is covered by $q$ drop boxes when $q \geq 2$.

If the optimal solution to the DBLP locates two or fewer drop boxes, the collection tour visiting the drop box locations is trivial. Thus, we assume that at least three drop boxes are be selected in the optimal solution. Under this assumption, we can formulate the DBLP using the following IP.

$$\begin{align*}
\text{min } z_1 &= \sum_{(i,j) \in E} c_{ij} x_{ij} + \sum_{j \in N} f_j y_j \\
\text{s.t. } r(v^0_w + v^1_w + \sum_{j \in N} a_{\mu} y_j) &\leq v^0_w + \sum_{j \in N} a_{\mu} y_j \quad \forall w \in W \\
\sum_{j \in N_w} y_j &\geq q \quad \forall w \in W \\
y_j &= 1 \quad \forall j \in T \\
\sum_{i \in N : (i,j) \in E} x_{ij} &= 2 y_j \quad \forall j \in N \\
\sum_{(i,j) : (i,j) \in E \text{ or } j \in S, i \in N \setminus S} x_{ij} &\geq 2 y_t \quad \forall S \subseteq N, 2 \leq |S| \leq |N| - 2, T \neq \emptyset, i \in S \\
\end{align*}$$

The objective (1) is to minimize the total cost of the drop box system. Constraint (2) requires that the value of the access function is at least $r$ for each $w \in W$. Constraint set (3) ensures that each $w \in W$ is covered by at least $q$ drop boxes within their respective covering set. Constraint set (4) ensures that a drop box is added at each location in $T$. Constraint sets (5) and (6) are used to determine the collection tour over the selected drop box locations using constraints originally introduced for the CTP (Gendreau et al., 1997). Constraint set (5) ensures that each selected drop box location is visited by the collection tour exactly once. Constraint set (6) introduces subtour elimination constraints. Note that these constraints differ from the subtour elimination constraints commonly seen in the TSP, since not all locations $N$ must be visited by the collection tour. The bound on the summation refers to the edges in $E$ such that the edge is incident to one node in $S$ and one in $N \setminus S$. Constraint sets (7) and (8) require the decision variables to be binary.

The Supplementary Online Materials file presents model properties. It notes that the DBLP is NP-hard and that some constraints can be removed for certain input parameters (see Section A.1). Additionally, the Supplementary Online Materials presents DBLP model variations DBLP for tailoring the model to various situations (see Section A.2).

4. Solution methods

4.1. Objective reformulation

Constraint sets (3)-(6) are similar to constraints that may be found in an IP for the CTP (Gendreau et al., 1997). However, the objective of the CTP only considers operational costs (Gendreau et al., 1997). Thus, it is desirable to reformulate objective $z_1$ to preserve properties of the CTP within the DBLP. We can then use components of solution methods for the CTP within solution methods for the DBLP, including an exact method based on lazy constraints (Section 4.2) and a heuristic method based on CTP and Group Steiner Tree problem heuristics (see Section A.3 in the Supplementary Online Materials).

We present a reformulation of $z_1$ to remove the use of $y_n$ variables. Note that constraints (5) enforce that for any drop box $n$ visited by a feasible tour, there must be exactly two drop boxes visited before or after $n$. Thus, we can reformulate $z_1$ as follows:

$$\begin{align*}
z_1 &= \sum_{i,j : (i,j) \in E : \text{ or } j \in S, i \in N \setminus S} c_{ij} x_{ij} + \sum_{j \in N} f_j y_j \\
&= \sum_{i,j : (i,j) \in E : \text{ or } j \in S, i \in N \setminus S} c_{ij} x_{ij} + \sum_{j \in N} f_j y_j \\
&= \sum_{i,j : (i,j) \in E} c_{ij} x_{ij} + \sum_{j \in N} f_j y_j \\
&= \sum_{i,j : (i,j) \in E} (c_{ij} + f_j / 2 + f_j / 2) x_{ij} \\
&= \sum_{i,j : (i,j) \in E} \tilde{c}_{ij} x_{ij} \\
\end{align*}$$

See election administrators likely have a fixed budget, but the amount allocated to managing the drop box system is likely not predetermined. Thus, we wish to minimize the proportion of the budget allocated to the drop box system.

$\tilde{c}_{ij}$When $q \geq 1$.

$\tilde{c}_{ij}$It can be easily checked whether two or fewer drop boxes are needed to satisfy the constraints of the model.
where $\hat{c}_{ij} := c_{ij} + f_i/2 + f_j/2$ for each $(i,j) \in E$. With this reformulation, $z_t$ takes the same form as the standard objective for the CTP.

### 4.2. Lazy constraint method

Branch and bound is one of the most common techniques used to solve IPs, and we employ it to solve the DBLP. However, constraint set (6) defines an exponential number of constraints, so we introduce a lazy constraint approach to solve the DBLP. First, we solve the DBLP without constraint set (6). Once an optimal solution is found, we determine if any of the constraints from constraint set (6) are violated. If so, we add in at least one violated constraint. Most modern optimization packages support the implementation of lazy constraints. The reformulation of the objective introduced in Section 4.1 can be used throughout the procedure, but it is not required.

We introduce a polynomial time algorithm, Algorithm 1, to find violated inequalities from constraint set (6) given an $x^* \in \{0,1\}^{|E|}$. The approach we take is adapted from an approach used for the TSP to account for the fact that not all potential drop box locations must be visited by the tour in the DBLP. Algorithm 1 first finds all subtours defined by $x^*$ (line 1). Each subtour that does not include all required locations $T$ (lines 2-4) must be associated with at least one violated constraint. For all $t$ locations $t$ visited by the subtour, we add the violated constraint (line 7).

**Algorithm 1 Lazy($x^*$)**

1: $H = \text{collection of subtours defined by } x^*$
2: for each subtour $h \in H$ do
3: \hspace{1cm} $\hat{S} = \text{drop box locations visited by } h$
4: \hspace{1cm} if $\hat{T} \setminus \hat{S} \neq \emptyset$ then
5: \hspace{1cm} \hspace{1cm} return $\sum_{(i,j) \in E \cap \hat{S}, j \in N \setminus \hat{S}} x_{ij} \geq 2y_t$ for each node $t \in \hat{S}$
6: \hspace{1cm} end if
7: end for

We comment on the correctness of Algorithm 1. Specifically, given an integer $x^* \in \{0,1\}^{|E|}$, Algorithm 1 finds a violated constraint from constraint set (6), if one exists. If a constraint is violated, there must exist a $S$ such that $\hat{S} \in N, 2 \leq |S| \leq |N| - 2, T \setminus \hat{S} \neq \emptyset$ and for some $t^* \in T \setminus \hat{S}$, $\sum_{(i,j) \in E \cap \hat{S}, j \in N \setminus \hat{S}} x_{ij} < 2y_{t^*}$. Since the left-hand side of the inequality is at least zero, $t^*$ must represent a selected drop box location ($y_{t^*} = 1$). Moreover, the feasibility of $x^*$ with regards to constraint set (5) implies that $\sum_{(i,j) \in E \cap \hat{S}, j \in N \setminus \hat{S}} x_{ij} = 0$. Thus, $t^*$ must be a member of some subtour visiting locations $\hat{S} \subseteq S$. The set $\hat{S}$ must contain at least three elements and can contain no more than $|N| - 3$ elements as a result of constraint set (5). Since $T \setminus \hat{S} \neq \emptyset$, it is also true that $T \setminus \hat{S} \neq \emptyset$. Thus, the existence of a $\hat{S}$ implies the existence of a $S$ whose elements form a subtour in $x^*$ such that $\hat{S} \subseteq N, 2 \leq |S| \leq |N| - 2, T \setminus S \neq \emptyset$ and $\sum_{(i,j) \in E \cap \hat{S}, j \in N \setminus S} x_{ij} < 2y_t$ for all $t \in \hat{S}$. Algorithm 1 identifies $\hat{S}$ and returns the corresponding constraint.

### 5. Case study

We construct a case study of Milwaukee, Wisconsin, USA to demonstrate the value of the DBLP and investigate the implications of optimal drop box infrastructure design. The City of Milwaukee is the most populous municipality in the state of Wisconsin and had approximately 315,483 registered voters prior to the 2020 General election (City of Milwaukee Open Data Portal, 2021). We let $W$ be defined by the census block groups of Milwaukee, WI Milwaukee County (2018), which are comprised of individuals located near each other who are typically of similar socioeconomic backgrounds. Figure 2(a) illustrates the different block group locations in Milwaukee and the estimated number of individuals of age 18 or older in each (United States Census Bureau, 2020).

During the 2020 elections, 15 drop boxes were placed throughout Milwaukee (Milwaukee Election Commission, 2020), illustrated in Figure 2(b). We use the DBLP to identify drop box locations assuming that these 15 were not already added to the voting system. This allows us to compare the DBLP to the decisions actually made by election officials during 2020. We let the potential drop box locations, $N$, be the locations of courthouses (4), election administrative buildings (2), fire stations (30), libraries (14), police stations (7), CVS pharmacies (7), and Walgreens pharmacies (29). Figure 2(c) illustrates the locations of the 93 potential drop box locations. We assume that the collection tour begins and ends at the Milwaukee City Hall, $s$. We do not require a drop box be located at any location other than City Hall, with $T = \{s\}$. The fixed cost of locating a drop box at court houses, fire stations, police stations, and City Hall is set at $6000 to reflect the cost of a drop box without the need of additional security measures. The fixed cost of locating a drop box at all other locations is set at $10,000 to reflect the cost of both a drop box and a security system.

According to the Milwaukee Election Commission, ballots were, at a minimum, collected daily by staff during the 2020 General election (Milwaukee Election Commission, 2020). This equates to approximately 50 times during the election. Based on this value, we assume that ballots are collected 50 times per year on average over the life of the drop boxes, which we assume to be 15 years. We further assume that each member of the bipartisan collection team has an opportunity cost of $40 per hour. This may not reflect the actual pay rate of poll workers or staff; rather, it is meant to represent the opportunity cost of other tasks not completed during that time. For example, staff could otherwise

---

11In reality, the number of collections depends on the election year. Also, the frequency of ballot collection may vary depending on model solutions, but this value is set to normalize the operational cost to the fixed cost of the drop boxes.
participate in additional security training, review compliance of submitted ballots, or conduct marketing to increase voter turnout. The cost of traveling between two drop boxes is determined using this pay rate and the estimated time needed to drive between the two locations, which is obtained from Bing Maps. We include the cost of gas and vehicle wear using the current federal mileage reimbursement of $0.56 per mile. The estimated mileage is calculated assuming an average travel speed of 30 mph. Lastly, we assume the collection costs increase by 2% each year.

The covering set of each location, \( N_w \), is constructed to include the locations that satisfy at least two of the following: the time to walk to the drop box is no more than 15 minutes; the time to drive to the drop box is no more than 15 minutes; the time to use public transit (i.e., city bus) to the drop box is no more than 30 minutes; or the road distance to the drop box is no more than 4 miles (e.g., reachable by bike or ride-share). By ensuring at least two conditions are met, there must be multiple transportation modalities that can be used to travel to a drop box in \( N_w \). Individuals without access to a private vehicle are thereby guaranteed to be able to reach a covering drop box using another mode of transportation.

We estimate the location of each block group centroid. Throughout the case study, we let \( q = 2 \), unless otherwise noted, so that model solutions mitigate the risk associated with the destruction of a drop box. Lastly, the parameters \( v^0_w, v^1_w, a_{nw} \) for the access function are instantiated as follows using a function that serves as a proxy based on historical voter turnout, transit durations obtained from Bing Maps, vehicle availability of individuals living in each block group (United States Census Bureau, 2019), and the work locations of individuals residing each block group (United States Census Bureau, 2021). The Supplementary Online Materials describe this function in detail (see Section A.5).

### 5.1. The DBLP and rules-of-thumb

In the absence of tools to support election planning, election administrators may use rules-of-thumb to select drop box locations. In this section, we demonstrate that the DBLP is able to identify drop box locations that outperform rules-of-thumb across multiple criteria. The findings support the use
of the DBLP during election planning. Table 2 presents the details of DBLP solutions for different values of $r$ obtained using the Gurobi 9.1 MIP solver. Computational studies were run using 64-bit Python 3.7.7 on an Intel® Core™ i5-7500 CPU with 16 GB of RAM. Each optimal solution was identified in less than 3600 seconds. We refer to the solutions identified by the DBLP as “DBLP $k$” where $k$ refers to the solution ID in Table 2.

During 2020, the Milwaukee Election Commission located drop boxes at the City Hall, the Election Commission warehouse, and 13 neighborhood-based public library branches (Milwaukee Election Commission, 2020). We begin by comparing these locations to those identified in DBLP 2, which also represents 15 drop boxes. Table 3 provides the values of multiple criteria for each drop box system. These criteria provide insight into the performance of each drop box system with respect to cost, access, and risk. Note that the 2020 policy is not a feasible solution to the DBLP, since it does not cover all voters (constraint set (3)).

The results in Table 3 suggests that the DBLP is able to identify drop box locations that perform better across multiple criteria compared with the rule-of-thumb approach used by election administrators in Milwaukee during the 2020 election. We find that with the same number of drop boxes, the DBLP is able to identify drop box locations that result in a lower fixed cost, operational cost, and total cost. Despite having a lower cost, all voters are covered by at least two drop boxes with DBLP 2, whereas the 2020 policy only covers 88.9% of voters twice. This gap also exists when voters do not have access to a vehicle (1.00 vs. 0.941). This means that DBLP 2 admits a higher level of equity of access to the drop box system while mitigating the risk associated with the possible destruction of a drop box. Moreover, the minimum access function value is higher (0.593 vs. 0.560) for DBLP 2. With a strict interpretation of the access function, the block group with the lowest turnout is predicted to have a turnout that is 3.3% higher if the DBLP 2 was implemented rather than the actual locations. We find that the average access function value is lower for DBLP 2 than the actual implementation; however, the difference is small (0.772 vs. 0.776).

In different situations, other rules-of-thumb may be used by election administrators. We compare the DBLP solutions to six other rules-of-thumb that could have potentially been used instead to demonstrate the value of the DBLP:

- Policy 1 Locate drop boxes at the election administrative buildings (2).
- Policy 2 Locate drop boxes at the election administrative buildings (2) and police stations (7).
- Policy 3 Locate drop boxes at the election administrative buildings (2) and libraries (14).
- Policy 4 Locate drop boxes at the election administrative buildings (2), police stations (7), and libraries (14).
- Policy 5 Locate drop boxes at the election administrative buildings (2) and fire stations (30).
- Policy 6 Locate drop boxes at the election administrative buildings (2), police stations (7), libraries (14), and fire stations (30).

These six policies yield six solutions for comparison that locate drop boxes at buildings that are well-distributed throughout the city. They are not intended to represent a comprehensive list of possible policies. Table 4 provides the values of multiple criteria (the same as in Table 3) for each rule-of-thumb and four DBLP solutions with a similar number of drop boxes. There are a few notable observations from Table 4. First, most rules-of-thumb are not feasible for the DBLP. Policy 6, which locates 53 drop boxes, is the only rule-of-thumb policy considered that guarantees that each $w \in W$ is covered by $q = 2$ drop box locations. Meanwhile, the DBLP finds feasible solutions with as few as 15 drop boxes. Second, the operational costs of the DBLP solutions are significantly lower than the corresponding rule-of-thumb solutions, due to the DBLP identifying short collection tours, leading to less work for the bipartisan team, particularly when there are many drop boxes. For example, the DBLP 8 solution (with 33 total drop boxes) has a operational cost of $14,501 as compared with an operational cost of $18,328 for Policy 5 (with 32 drop boxes). Third, the DBLP identifies drop box locations that are consistently better across multiple criteria than rules-of-thumb with a similar number of criteria.

| Criteria                                                                 | 2020          | DBLP 2         |
|--------------------------------------------------------------------------|---------------|----------------|
| Number of drop boxes                                                     | 15            | 15             |
| Fixed cost ($/year)                                                      | 9733          | 7333           |
| Operational cost ($/year)                                                | 10,566        | 10,479         |
| Total cost ($/year)$                                      | 20,300        | 17,813         |
| Fraction of voters covered by one drop box (population weighted)        | 0.995         | 1.00           |
| Fraction of voters covered by two drop boxes (population weighted)†       | 0.889         | 1.00           |
| Fraction of voters without access to a car covered by at least two drop boxes by non-driving transit (population weighted)‡   | 0.941         | 1.00           |
| Minimum access function value§                                           | 0.560         | 0.593          |
| Average access function value (population weighted)                      | 0.776         | 0.772          |
| Maximum road distance to closest drop box                               | 7.634         | 6.311          |
| Maximum road distance to third closest drop box                          | 10.55         | 9.978          |
| Average road distance to closest drop box (population weighted)          | 1.601         | 1.679          |
| Average road distance to closest three drop boxes (population weighted)  | 2.723         | 2.486          |

*Required by the DBLP.
†Required by constraint set (3) of the DBLP.
‡Required by constraint set (3) given our method of instantiating $N_w$ for each $w \in W$.
§Modeled using constraint set (2) in DBLP.
suggest there is a substantial trade-off between the cost of the drop box system and the minimum access function value. However, the marginal increase in cost to achieve an increase in the minimum access function value is not constant. From DBLP solution 0 to DBLP solution 1 the average cost of a 0.01 increase of the minimum access function value is $186.84 per year (with 186.84 ≈ $17,313−$17,414). From solutions 3 to 4 the average cost of a 0.01 increase of the minimum access function value is $50,738 per year. This highlights the importance of considering the access function within the DBLP. When a low-cost solution is desirable, an appropriate value for $r$ allows the DBLP to identify drop box locations that admit a much larger minimum access function value for a relatively low increase in cost (e.g., solutions 1–4). When drop boxes that admit a large minimum access function value are desirable, it is critical to appropriately set $r$, since a small change in $r$ can lead to solutions of substantially different cost (e.g., DBLP solutions 8–10 in Table 2).

We next consider the trade-off between equitable access to the drop box system and equitable access to all voting pathways. Figure 3 plots the cost and minimum access function value of multiple optimal solutions when $q$ is zero (−−−−), one (−−−−−−), or two (−−−−−−−−−−−−−−) with the latter corresponding to the solutions presented Table 2. When $q=0$, the DBLP is able to identify drop box locations that substantially increase the minimum access function value for a relatively small cost. This suggests that there are cost-effective, equitable drop box locations, even when election officials cannot afford to cover each voter with one or two drop boxes. In general, equitable access to the drop box system and equitable access to all voting pathways are aligned so that access is improved. However, the difference between the curves corresponding to $q=0$ (−−−−) and $q=1$ (−−−−−−) represents the cost of ensuring equitable access to the drop box system. In some cases, this cost can be substantial ($\sim$ $6767$ per year). This demonstrates the trade-off between selecting drop boxes that ensure all voters have access to the drop box system or using the drop boxes to increase the

### 5.2. Drop box trade-offs

In this section, we further investigate DBLP solutions and explore the trade-offs between criteria within the drop box voting system. We begin by discussing the trade-off between cost and equity of access to all voting pathways (i.e., the minimum access function value). The solutions in Table 2

**Figure 3.** Solutions using different values of $q$. drop boxes. The DBLP solutions are generally more equitable with lower maximum road distances to the closest and third closest drop boxes. Notably, the DBLP solutions have higher minimum access function values than their corresponding rules-of-thumb. The DBLP 6 solution with 23 drop boxes has a minimum access function value of 0.629, which is higher than those of all the rule-of-thumb solutions. One exception where the rules-of-thumb perform well is the average road distance to the closest drop box (population weighted). Policy 3, for example, has an average distance of 1.550 miles, which is lower than the 1.689 miles of the DBLP 3 solution with the same number of drop boxes. In sum, these observations shed light on how the DBLP drop box locations add value as compared with alternative approaches.

**Table 4.** A comparison of rule-of-thumb and DBLP policies across multiple criteria.

| Criteria | Policy 1 | Policy 2 | Policy 3 | Policy 4 | Policy 5 | Policy 6 | DBLP 3 | DBLP 6 | DBLP 8 | DBLP 10 |
|----------|----------|----------|----------|----------|----------|----------|--------|--------|--------|--------|
| Number of drop boxes | 2 | 9 | 16 | 23 | 32 | 53 | 16 | 23 | 33 | 52 |
| Fixed cost ($/year) | 1067 | 3867 | 10,400 | 13,200 | 13,067 | 25,200 | 8267 | 12,133 | 17,200 | 27,200 |
| Operational cost ($/year) | 1535 | 7233 | 10,616 | 11,954 | 18,328 | 21,129 | 10,271 | 11,466 | 14,501 | 19,823 |
| Total cost ($/year) | 2602 | 11,100 | 21,016 | 25,154 | 31,395 | 46,329 | 23,599 | 31,701 | 47,023 | 60,329 |
| Minimum access function value of multiple optimal solutions when | $q=0$ | $q=2$ | $q=3$ | $q=4$ | $q=5$ | $q=6$ | $q=7$ | $q=8$ | $q=9$ | $q=10$ |
| Average access function value (population weighted) | 0.542 | 0.558 | 0.568 | 0.582 | 0.591 | 0.623 | 0.601 | 0.629 | 0.645 | 0.661 |
| Average road distance to closest drop box (population weighted) | 5.829 | 2.130 | 1.550 | 1.469 | 1.062 | 0.917 | 1.689 | 1.517 | 1.386 | 1.105 |
| Fraction of voters covered by 1 drop box (population weighted) | 0.362 | 0.810 | 0.924 | 0.973 | 0.997 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Fraction of voters covered by 2 drop boxes (population weighted) | 0.467 | 0.920 | 0.963 | 0.981 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Minimum access function value of multiple optimal solutions when | $q=0$ | $q=2$ | $q=3$ | $q=4$ | $q=5$ | $q=6$ | $q=7$ | $q=8$ | $q=9$ | $q=10$ |
| Average access function value (population weighted) | 0.760 | 0.770 | 0.777 | 0.786 | 0.790 | 0.809 | 0.773 | 0.781 | 0.788 | 0.806 |
| Average road distance to closest drop box (population weighted) | 5.829 | 2.130 | 1.550 | 1.469 | 1.062 | 0.917 | 1.689 | 1.517 | 1.386 | 1.105 |
| Fraction of voters covered by 1 drop box (population weighted) | 0.362 | 0.810 | 0.924 | 0.973 | 0.997 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Fraction of voters covered by 2 drop boxes (population weighted) | 0.467 | 0.920 | 0.963 | 0.981 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

| Image 81x347 to 267x529 |

| Image 316x219 to 333x224 |

Note:**Objective of the DBLP.**

*Required by constraint set (3) of the DBLP.

*Required by constraint set (3) given our method of instantiating $N_w$ for each $w \in W$.

*Modeled using constraint set (2) in DBLP.
access function value by “filling in the gaps” of the in-person voting system. We also find a substantial difference between the curves corresponding to $q = 1$ (---) and $q = 2$ (-----), particularly when a low-cost solution is desired. This suggests that the cost of mitigating risks associated with the destruction of drop boxes through infrastructure design is relatively high and may not be cost-effective. Instead, it may be more cost-effective to respond to an adverse event after it occurs, since the likelihood of this risk occurring is low (Scala et al., 2022).

Rather than changing $q$, we can also relax coverage by defining the coverage sets $N_w$ using a larger time threshold. When covering sets are defined by a longer time threshold, the drop boxes are allowed to be located further away from the voters while still meeting the coverage constraints defined in constraint set (3). A larger time threshold may increase the inequity of access to the drop box infrastructure within the resulting solutions, since a census block group may be further from both covering drop boxes when compared with other census block groups. However, the access function continues to evaluate the effect of the drop box locations on voter turnout when all voting pathways are considered. Figure 4 illustrates the cost and minimum access function value of solutions to the DLBP using covering sets $N_w$ obtained using the same procedure as before, except the time threshold are multiplied by factors of 0.9 (-----), 1.0 (-----), 1.1 (-----), 1.2 (-----), and 1.3 (-----) when $q = 2$. A factor of 1.0 corresponds to the covering sets used to obtain the solutions discussed in Table 2 and Figure 3. We find the effects of changing the covering sets to be similar to the effect of changing $q$. Covering sets defined by a larger factor result in solutions with a larger minimum access function value for the same cost.

We explore solutions DBLP 0, 3, and 6 of Table 2 in more detail. Figure 5 illustrates the selected drop box locations and the collection tour visiting these drop boxes overlaid on a map of Milwaukee. Each black circle indicates a selected drop box location, and the blue lines describe the order in which the drop boxes are visited on the collection tour (not the actual roads driven). The color of each block group indicates the access function value of each block group with red reflecting relatively a low value and green reflecting a relatively high value. When cost is minimized (Solution 0), drop boxes are well-spaced in order to cover each block group twice, but relatively few drop boxes are placed to reduce cost. Solution 0 is notably different from the locations selected during 2020, Figure 2(b), in the northern and southern areas of the city despite locating the same number of drop boxes. The DBLP selects additional drop box locations in the north and southern part of the city, which would otherwise have have a relatively low level of access to the voting infrastructure,
indicated by the dark red in Figure 5(a). Additional locations are not selected in the south, since those voters have relatively high access to the multiple voting pathways, indicated by the dark green in Figure 5(a).

We study the performance of the DBLP heuristic compared with the lazy constraint approach in the Supplementary Online Materials (see Section A.6). These additional computational studies suggest that the proposed heuristic method approximates the Pareto frontier between cost and the minimum access function value well and does so quickly.

6. Conclusion

In this article, we introduce a structured and transparent approach to support the planning of ballot drop box voting systems, particularly for USA voting systems. We do so by formalizing the DBLP that identifies drop box locations that minimize cost while ensuring voters have access to the drop box system and drop box risks are mitigated. Using a real-world case study, we demonstrate that the DBLP identifies drop box locations that consistently outperform rules-of-thumb across multiple criteria. We also find that the trade-off between criteria is non-trivial and requires careful consideration.

Our research suggests that optimization is an important tool for designing the drop box infrastructure. Simple guidance for designing drop box systems, such as locating one drop box per 15,000 registered voters, or other rules-of-thumb may be overly-simplistic and can cause election administrators to overlook important criteria. Strategic drop box locations can reduce the “cost” of voting while ensuring that all voters have equitable access to the drop box system. Equity is increasingly becoming a legal requirement for locating drop boxes (Michigan state, 2022). Future research can utilize the DBLP to answer additional drop box policy questions to support the drafting of legislation surrounding the use of drop boxes.

We introduce a lazy constraint approach to solve the DBLP to optimality. Computational experiments show that a single optimal solution to the DBLP can be found relatively quickly using this approach within a state-of-the-art MIP solver for moderately sized problem instances. However, finding multiple solutions to the DBLP that perform well across multiple criteria can lead to computational times that are unreasonable in practice. This motivates the introduction of a heuristic for the DBLP that quickly identifies near-optimal solutions. Initial attempts at reducing the computational time needed to identify optimal solutions using cutting planes originally introduced for the CTP (Gendreau et al., 1997) proved unfruitful. Future research into the theory of the DBLP is needed to reduce solution times for exact methods.

The DBLP is intended to be a component of a larger suite of tools for supporting election administrators understand, assess, and ultimately design different facets of the voting infrastructure. Ideally, the DBLP and other operations research tools will eventually be integrated into an online platform designed to support election administrators in all aspect of elections planning. There is a substantial opportunity for the operations research community to support election planning by appropriately modeling voting systems and voting infrastructure. Future research is needed to understand the temporal aspects of risk, particularly in the absentee voting process, and determine best practices for mitigating against malicious and non-malicious attacks. The DBLP and future models can then be incorporated into a comprehensive tool to support election officials in designing the election infrastructure in a way that increases voter turnout. A key challenge within this space is the need to understand and incorporate models that describe how voters freely select from multiple voting pathways once the infrastructure is set. Voter choices ultimately determine the performance of the voting system.

Acknowledgments

The authors would also like to thank the anonymous reviewers, whose suggestions for improvement led to a substantially improved manuscript.

Notes on contributors

Adam Schmidt works at Johns Hopkins University Applied Physics Laboratory and is interested in protecting critical infrastructure and healthcare systems engineering. He earned his PhD in industrial engineering in 2022 from the University of Wisconsin-Madison.

Laura Albert is a Professor and the David Gustafson Department Chair of Industrial & Systems Engineering at the University of Wisconsin-Madison. Her research interests are in the areas of optimization and analytics with application to homeland security, emergency response, and public sector problems. She is a Fellow of the American Association for the Advancement of Science (AAAS) and the Institute of Industrial and Systems Engineers (IISE) Fellow Award. She is the author of the blog “Punk Rock Operations Research.”

References

Allen, T. and Bernshteyn, M. (2006, August) Mitigating voter waiting times. CHANCE, 19, 25–34.

Aloulou, F. (2018, November) The Application of Discrete Choice Models in Transport, IntechOpen, London.

Baldacci, R., Boschetti, M., Maniezzo, V. and Zamboni, M. (2005, January) Scatter search methods for the covering tour problem, in R. Sharda, S. Voß, C. Rego, B. Alidaee (eds.), Metaheuristic Optimization via Memory and Evolution, vol. 30, Springer, Boston, MA, pp. 59–91.

Burden, R. (2021) The experiences of municipal clerks and the electorate in the November 2020 general election in Wisconsin. Technical report, University of Wisconsin - Madison, Madison, WI.

Cantoni, E. (2020, January) A precinct too far: Turnout and voting costs. American Economic Journal: Applied Economics 12(1), 61–85.

Church, R. and ReVelle, C. (1974, December) The maximal covering location problem. Papers of the Regional Science Association, 32(1), 101–118.

City of Milwaukee Open Data Portal (2021) 2020 General Election - City of Milwaukee canvassed election results. Report, Milwaukee, WI.

Collingwood, L., McGuire, W., Gonzalez O’Brien, B., Baird, K. and Hampson, S. (2018, February) Do drop boxes improve voter turnout? Evidence from King County, Washington. Election Law Journal: Rules, Politics, and Policy, 17(1), 58–72.

Corasaniti, N., Shear, M.D. and Gabriel, T. (2020, August) Postal crisis has states looking for alternatives to mail-in ballots. The New York Times, https://www.nytimes.com/2020/08/17/us/politics/postal-service-voting.html (accessed 16 March 2021).
Davendra, D. (2010) *Traveling Salesman Problem: Theory and Applications*. IntechOpen, London.

Downs, A. (1957) *An Economic Theory of Democracy*. Harper and Brothers, New York.

EAC Advisory Board and Standards Board (2009, December) *Elections operations assessment*. Threat trees and matrices and threat instance risk analyzer (TIRA). Technical report, University of South Alabama.

Edelstein, W.A. and Edelstein, A.D. (2010) *Queuing and elections*. Long lines, dres and paper ballots, in *2010 International Conference on Electronic Voting Technology/Workshop on Trustworthy Elections*, USENIX Association, Washington, DC, pp. 1–8.

Fitzsimmons, Z. and Lev, O. (2020) *Selecting voting locations for fun and profit*, in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence* (IJCAI-20), Association for Computing Machinery, Yokohama, Japan, pp. 224–230.

Gendreau, M., Laporte, G. and Semet, F. (1997) *The covering tour problem*. *Operations Research*, 45(4), 568–576.

Glasgow, G. and Alvarez, R.M. (2008, August) *Discrete Choice Methods*. Oxford University Press, Oxford.

Greene, S. and Ueyama, K. (2015, April) *Vote-by-mail rates more than double since 2000*. Article, Pew Charitable Trusts, Philadelphia, PA.

Haspel, M. and Knotts, H.G. (2005, May) *Location, location, location: Precinct placement and the costs of voting*. *The Journal of Politics*, 67(2), 560–573.

Highton, B. (2006) *Long lines, voting machine availability, and turnout: The case of Franklin County, Ohio in the 2004 presidential election*. PS: Political Science and Politics, 39(1), 47–53.

Highton, B. (2006) *Long lines, voting machine availability, and turnout: The case of Franklin County, Ohio in the 2004 presidential election*. PS: Political Science and Politics, 39(1), 47–53.

Joint COVID Working Group (2020) *Ballot Drop Box* (1 ed.), 245 Murray Lane Washington, D.C. 20528-0380: Cybersecurity and Infrastructure Security Agency (CISA) Elections Infrastructure Government Coordinating Council and Sector Coordinating Council. https://static1.squarespace.com/static/5a6665c98017db2b60bc22084/t/5e8f4d717ce5e7ee2d8cb81586447064805/Ballot_Drop-Box_final.pdf.

Mavrotas, G. (2009) *Effective implementation of the ε-constraint method in multi-objective mathematical programming problems*. *Applied Mathematics and Computation*, 213(2), 455–465.

McGuire, W., O’Brien, B.G., Baird, K., Corbett, B. and Collingwood, L. (2020, September) Does distance matter? Evaluating the impact of drop boxes on voter turnout. *Social Science Quarterly*, 101(5), 1789–1809.

Michigan state (2022) *Ballot proposition 2 of 2022*. Proposition, Michigan House of Representatives. https://www.house.mi.gov/hfa/PDF/Alpha/Ballot_Proposal_2_of_2022.pdf (accessed 8 November 2022).

Milwaukee County (2018) Census blocks, Milwaukee County. Report, City of Milwaukee, Milwaukee, WI.

Milwaukee Election Commission (2020, September) *Absentee ballot drop boxes available at 15 locations across the city, including 13 Milwaukee Public Library branches, Milwaukee, WI*. https://city.milwaukee.gov/mayorbarrett/News/2020-News/Absentee-Ballot-Drop-Boxes-Available-at-15-Locations-across-the-City-Including-13-Milwaukee-Public (accessed 21 February 2022).

MIT Election Data + Science Lab (2021, March) Voting by mail and absentee voting. Report, MIT, Cambridge, MA.

MIT Election Data & Science Lab (2022) *Elections performance index*. Report, Massachusetts Institute of Technology, Cambridge, MA.

Murakami, K. (2018) *A generalized model and a heuristic algorithm for the large-scale covering tour problem*. *RAIRO-Operations Research*, 52(2), 577–594.

National Conference of State Legislatures (2022) Voting outside the polling place: Absentee, all-mail and other voting at home options. Report, National Conference of State Legislatures, Washington, DC.

Pew Research Center (2020a) Sharp divisions on vote counts, as Biden gets high marks for his post-election conduct. Report, Pew Research Center, Washington, DC.

Pew Research Center (2020b, November) *The 2020 voting experience: Coronavirus, mail concerns factored into deciding how to vote*, Washington, DC. https://www.pewresearch.org/politics/2020/11/20/the-voting-experience-in-2020/ (accessed 10 January 2022).

Scalia, N.M., Goethals, P., Dehlinger, J., Mezgebe, Y., Jilcha, B. and Bloomquist, I. (2022) *Evaluating mail-in security for electoral processes using attack trees*. *Risk Analysis*, 42(10), 2327–2343.

Schafer, B. and Gammann, B. (2020, September) *RI board of elections hopes to install security cameras at mail ballot drop boxes*. WPRI.com (accessed 16 March 2021).

Scherer, Z. (2021, April) Majority of voters used nontraditional methods to cast ballots in 2020. Report, U.S. Census Bureau, Washington, DC.

Schmidt, A. and Albert, L.A. (2021, October) *Designing pandemic-resilient voting systems*. *Socio-Economic Planning Sciences*, 80, 101174.

Sigelman, L. and Berry, W.D. (1982, December) *Cost and the calculus of voting*. *Political Behavior*, 4(4), 419–428.

Spencer, D.M. and Markovits, Z.S. (2010) *Long lines at polling stations? Observations from an election day field study*. *Election Law Journal*, 9(1), 3–17.

Stewart III, C. (2017) *2016 Survey of the performance of American elections*. Harvard Dataverse data set.

Stewart III, C. and Ansolabehere, S. (2015) *Waiting to vote*. *Election Law Journal*, 14(1), 47–53.

Toregas, C., Swain, R., ReVelle, C. and Bergman, L. (1971) *The location of emergency service facilities*. *Operations Research*, 19(6), 1363–1373.

Train, K. (2009) *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, UK.

United States Census Bureau (2019) *Means of transportation to work by selected characteristics*. Data table, U.S. Census Bureau, Washington, DC.

United States Census Bureau (2020) *Race for the population 18 years and over*. Data table, U.S. Census Bureau, Washington, DC.

United States Census Bureau (2021) *Work destination analysis by census tracts*. Census data, U.S. Census Bureau, Washington, DC.

Vargas, L., Jozefowiez, N. and Ngueveu, S.U. (2017) *A dynamic profit-maximizing model for corrupt elections*. *Journal of Heuristics*, 23(1), 53–80.

Vasilogambros, M. (2020, December) *Lawmakers push to preserve pandemic voting access*. Folsom, CA, Article on govtech.com (accessed 17 March 2021).

Wang, X., Yang, M. and Fry, M.J. (2015) *Efficiency and equity trade-offs in voting machine allocation problems*. *Journal of the Operational Research Society*, 66(8), 1363–1369.

Yang, M., Fry, M.J. and Kelton, W.D. (2009) *Are all voting queues created equal?* in *Proceedings of the 2009 Winter Simulation Conference* (WSC), IEEE Press, Picataway, NJ, pp. 3140–3149.