Districting Problems

2AMS50, Optimization for Data Science, Q3

Group 10

Full Name	Student ID
Luca Mainardi	2014602
Laurynas Jagutis	2037637
Ian van de Wetering	1009805
Błażej Nowak	1617303
Thomas Warnier	1423495

Eindhoven, March 28, 2024

Contents

1 | Introduction

The problem of political redistricting originated when Elbrige Gerry, governor of Massachusetts, signed a Democratic Party-friendly redistricting plan into law in 1812, which is now known as gerrymandering. Since then, the primary objective of the problem of political districting has been to divide geographical regions into electoral districts with fair representation and an equitable distribution of political power among voters. The problem in its current form presents three main constraints for districts:

- **Size:** equal population distribution among districts maintains fairness and prevents disparities in political influence, ensuring equitable representation for all.
- **Contiguity:** requires all district parts to be connected, fostering community cohesion and enabling effective representation by a single representative.
- **Compactness:** shapes districts to prevent gerrymandering, promoting fairness and transparency in representation by ensuring easily identifiable boundaries.

In order to find optimal solutions for the political districting problem, integer programming can be used because it allows for the formulation of constraints and objectives in mathematical terms. By representing districting decisions as binary variables, it enables the exploration of optimal solutions that satisfy contiguity, compactness, and size constraints. A model of this type is proposed in this report, as well as an alternative solution that uses metaheuristics, specifically simulated annealing.

To analyze our algorithms, we used a dataset consisting of 50 US states, where each state is divided into counties. We were then able to analyze the optimal solutions for the states selected for our experiments, to verify their correctness.

2 | Literature Study

One of the first known uses of computer optimization for the redistricting problem has been presented by Hess et al. (1965) [1]. Within this paper, they introduce a heuristic model based on the k-median warehouse-location problem. This method uses a transportation algorithm to assign populations equally to district centers while minimizing the total cost or sum of squared distances from each person to the district center. This is an iterative algorithm and for each iteration the district must be within the legislative district. Within this paper, the notion of contiguity and compactness also get their first introduction.

Another important method was presented in Garfinkel and Nemhauser (1970) [2], who developed the set partition model, where a binary variable is used for each possible district, where the goal is to select k such that each territory of the state is covered exactly once. Anuj (1998) [3] have proposed solutions based on this formulation. The main drawback of this model is that the set of possible districts grows exponentially. This means that the solution is possible using a selected subset of district variables or by introducing variables on the fly via column generation.

Other variations for integer programming are formulated by Shirabe (2009) [4]. Three models are formulated, where each model optimizes for one of the constraints; contiguity, compactness and size. They found that the first and second constraint are the easiest to solve, while the last constraint takes significantly longer. Validi (2022) [5] focuses on the best method to impose the contiguity constraint in the context of the Hess model, since previous research had difficulty to explicitly ensure the contiguity constraint. They studied four algorithms, two flow-based algorithms (SHIR, MCF) and two cut-based algorithms (CUT, LCUT) and analyzed their ability to ensure the contiguity constraint and the execution times for each algorithm.

Since every variant of redistricting is NP-hard [6], many researchers have studied the use of heuristics. Greedy construction was proposed by Vickrey (1961) [7]. The idea for greedy construction is to iteratively extend the districts. Whenever the goal is to improve an existing distribution, we can use the local search heuristic implemented by King (2012, 2014, 2018) [8][9][10]. Some metaheuristics were suggested, such as simulated annealing and tabu search [11][12][13]. More recently, Cho and Liu (2018) [14] and Fifield (2020) [15] proposed Markov chain Monte Carlo (MCMC) methods to generate redistricting plans. These methods are very similar to local search, as they move from a feasible solution to a neighboring feasible solution. Many of these methods are also applied outside the context of political redistributing, like in the context of deregulated electricity markets districting [16] and artificial bee colony districting [17].

In addition to the computational and algorithmic challenges inherent in political redistricting, the work of Rios-Mercado (2020) [18] introduces an essential perspective on the socio-political dimensions of the problem. The design of political districts must not only aim for equitable population distribution and geographical compactness but also consider administrative boundaries, social-demographic characteristics, and the historical and current contexts of minority representation as mentioned in the Introduction 1. This necessitates the use of sophisticated tools such as Geographic Information Systems (GIS) to handle spatial data effectively, allowing for a nuanced analysis that can accommodate these diverse criteria. Such considerations underscore the complexity of districting as a problem that intersects with issues of fairness, representation, and social justice, highlighting the importance of incorporating a wide range of factors into the optimization process.

3 | Problem Formulation

3.1 | Mathematical Formulation

For the mathematical formulation of the problem it is useful to define the adjacency graph $(G = (V, E))$, in which each vertex $v \in V$ represents a territory (for example a county or a census tract) and in which there is a branch $u, v \in E$ which connects the vertices u and v when the relative territories are neighboring. Other necessary data are the number k of districts that must be formed, the population p_v of each territory $v \in V$ and the distance d_{ij} between territory i and territory j, fundamental for defining the compactness of the districting plan, which will appear in the objective function in the model. The minimum and maximum population allowed in a district (L and U) must also be defined. Currently in the USA there is no threshold for population deviation above which a district plan will necessarily be accepted. For example, a population deviation of only 19 people (0.0029%) was declared unconstitutional in 2002 by a federal district court in Pennsylvania, while population deviations of up to 1% may be permitted if there is a convincing justification (West Virginia kept all of its counties intact at the price of a 0.79% population deviation [19]).

The IP formulation for the political districting problem was introduced by Hess et al (1965) [1]. In this formulation n^2 binary variables are used:

$$x_{ij} = \begin{cases} 1 & \text{if vertex } i \text{ is assigned to (the district centered at) vertex } j \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

The formulation of the problem is as follows:

$$\begin{aligned}
\min \quad & \sum_{i \in V} \sum_{j \in V} w_{ij} x_{ij} && && \text{(1a)} \\
\text{s.t.} \quad & \sum_{j \in V} x_{ij} = 1 && \forall i \in V && \text{(1b)} \\
& \sum_{j \in V} x_{jj} = k && && \text{(1c)} \\
& L x_{ij} \leq \sum_{i \in V} p_i x_{ij} \leq U x_{jj} && \forall j \in V && \text{(1d)} \\
& x_{ij} \leq x_{jj} && \forall i, j \in V && \text{(1e)} \\
& x_{ij} \in \{0, 1\} && \forall i, j \in V. && \text{(1f)}
\end{aligned}$$

Constraints (1b) ensure that each vertex is assigned to a district. Constraint (1c) ensures that k districts are chosen. Constraints (1d) ensure that the population of each district lies between L and U. Constraints (1e), although not originally included by Hess et al.[1], are added for strength [20]. They impose that if a territory i is assigned to a district centered in j ($x_{ij} = 1$), then territory j must necessarily be assigned to the district of which it is the center ($x_{jj} = 1$).

The objective function considers the moment of inertia, in which there are penalties w_{ij}, defined as follows:

$$w_{ij} := p_i d_{ij}^2$$

Thus, we are trying to minimize the total cost of districting, where the cost of assigning a vertex to a district is proportional to the population of the vertex multiplied by the square of the distance between the vertex and the center of the district. This implies that both population density and geographic distance are being considered when subdividing districts.

The Hess model described above was used as the foundation for our proposed solution. The feasible region of its LP relaxation, which we denote by P_{HESS}, is defined as follows.

$$P_{HESS} := \{x \in \mathbb{R}^{n \times n} \text{ s.t. } x \text{ satisfies constraints (1b), (1c), (1d), (1e)}\}.$$

As can be seen, none of the model constraints explicitly impose contiguity. Accordingly, Hess et al. [1] had to manually adjust their solutions to make them contiguous. However, due to the model's compactness-seeking objective, it tends to generate contiguous or nearly contiguous districts.

Oehrlein and Haunert [21] proposed a flow-based formulation, adapted from Shirabe [4] which we used to impose contiguity on the Hess model. This formulation uses the following flow variables:

f_{ij}^v = the amount of flow, originating at district center v, that is sent across edge (i, j).

First, we introduce the "bidirected" version of the contiguity graph which we denote by $D = (V, A)$. This directed graph D is obtained from $G = (V, E)$ by replacing each undirected edge $i, j \in E$ by its directed counterparts (i, j) and (j, i). Thus, $|A| = 2|E|$. The set of edges pointing away from vertex i is denoted by $\delta^+(i)$, and the set of edges pointing towards vertex j is denoted by $\delta^-(i)$.

The formulation is as follows:

$$
\begin{aligned}
& x \in P_{HESS} & & \text{(2a)}\\
& f^j(\delta^-(i)) - f^j(\delta^+(i)) = x_{ij} & \forall i \in V \backslash \{j\}, \forall j \in V & \quad \text{(2b)}\\
& f^j(\delta^-(i)) \leq (n-1)x_{ij} & \forall i \in V \backslash \{j\}, \forall j \in V & \quad \text{(2c)}\\
& f^j(\delta^-(j)) = 0 & \forall j \in V & \quad \text{(2d)}\\
& f_{ij}^v \geq 0 & \forall (i,j) \in A, \forall v \in V & \quad \text{(2e)}\\
& x_{ij} \in \{0, 1\} & \forall i, j \in V. & \quad \text{(2f)}
\end{aligned}
$$

Constraints (2b) ensure that if vertex i is assigned to center j, then i consumes one unit of flow of type j; otherwise, it consumes none. Constraints (2c) ensure that vertex i can receive flow of type j only if i is assigned to center j. Constraints (2d) prevent flow circulations and constraints (2e) ensures that the amount of flow is non-negative.

4 | Solution Method

4.1 | Exact Solution

The exact solution to the districting problem is implemented using a mathematical optimization model, specifically formulated to balance compactness, contiguity, and population equality across districts in a given state. Employing the Gurobi optimization tool, the model defines decision variables to represent the assignment of geographical nodes to districts, ensuring each district is formed by contiguous nodes and the population variance across districts does not exceed a predefined threshold. The objective function aims to minimize the sum of squared distances between nodes, weighted by their populations, to create compact districts. Constraints are designed to enforce contiguity, ensure each node belongs to exactly one district, and maintain population thresholds. The model's effectiveness resides in its detailed formulation and the robust optimization capabilities of Gurobi, which together enable the precise identification of an optimal districting. This approach, while computationally demanding, leverages advanced mathematical and computational techniques to address a complex problem often encountered in geographical planning, showcasing the power of optimization software in solving real-world challenges exactly.

4.2 | Heuristics Solution

The solution combines a greedy heuristic method with simulated annealing to solve a districting problem, aiming for districts that are contiguous, balanced in population, and compact. Initially, the method sets up adjacent districts by allocating geographical nodes in such a way that each district comprises of neighboring nodes. It is achieved through a heuristic that expands districts starting from seed nodes chosen at random. This creates an initial solution that, while not necessarily optimal, serves as a starting point. This solution makes sure that all of the provinces within the districts are connected, as this is required for the next step, but it does not take into account the population constraints.

Next, the method employs a simulated annealing algorithm, a probabilistic technique for approximating the global optimum of a given function. It iterates through neighboring solutions by swapping border nodes between districts to improve the balance of population across districts, guided by an objective function that penalizes deviations from ideal population levels and rewards compactness. The algorithm accepts a new solution if it improves the objective function or, probabilistically, if it introduces diversity into the search process, which helps to escape local optima.

A neighbor solution is considered valid if it maintains or improves the population balance within a predefined threshold and ensures all districts remain contiguous. Contiguity is verified through a depth-first search algorithm, ensuring that each district is a single, unbroken territory. The balance of population is assessed against an ideal target, with penalties for deviations, encouraging solutions where districts are as equal in population as possible.

The objective function combines these considerations, weighting population balance and compactness, to guide the search towards solutions that are both balanced and efficient in terms of the geographic distribution of districts. The algorithm's temperature parameter gradually decreases, reducing the amount of exploration done, thus focusing the search on the most promising areas of the solution space.

This method leverages the strengths of heuristics and simulated annealing techniques to explore and refine solutions, ensuring that the final districts meet the criteria. The main drawback of this method is the fact that it is hardly dependent on the initial distribution of regions among the districts. Due to that the final solution it presents not always meets the previously defined population constraints.

5 | Results

Due to the limited computational power at our disposal, solving the political redistricting problem with data at the census tract level was unsolvable in most of its instances using both models we developed. Therefore, we decided to use only county-level data.

In a redistricting contest organized by Ohio reformers [22], the value of the ideal population $p := \frac{1}{k}\sum_{i\in V} p_i$ of a district was calculated and a deviation of 1% was allowed, setting $L := 0.995(p)$ and $U := 1.005(p)$, appropriately rounded to an integer.

However, using this threshold most cases at the county level are not possible, as also demonstrated by Validi et al. (2021) [5]. To illustrate an example where a 1% threshold is not feasible, consider Texas, the population of Dallas County is $p_v = 2368139$ residents, far exceeding the imposed population limit of $U = 701980$. Texas is therefore clearly unworkable at the county level. Therefore, we have decided to employ several varying threshold values (10%, 20%, 50%) to assess their impact on model feasibility.

Several states exhibit a type of blatant unachievability like Texas, where the population of a single vertex is greater than U even at a 50% threshold. We excluded these states from our experiments, except WA, which is required by assignment delivery. Even the trivial cases whose number of districts k is 1 have been excluded, except AK, also in this case as it is explicitly required by the assignment. From the experiments of Validi et al. [5] it also turns out that CO, NH, SC and OR are infeasible when contiguity is imposed. These states are therefore also excluded from our analyses. From the remaining states, we decided to select 7: AK, IA, ID, KS, MS, RI, WA.

5.1 | Exact Solution

Table 5.1: Redistricting results, with exact solution

State	n	k	Threshold	Moment of Inertia	District	Population
MS	82	4	0.1	15623364045.0	1	759102
					2	708866
					3	717670
					4	775641
			0.2	15614471350.0	1	790087
					2	705066
					3	690485
					4	775641
			0.5	15072865905.0	1	729877
					2	865237
					3	558300
					4	807865
AK	177	1	0.1	163327013658.0	0	733391
			0.2	163327013658.0	0	733391
			0.5	163327013658.0	0	733391
WA	39	10	0.1	infeasible		
			0.2	infeasible		
			0.5	infeasible		
IA	99	4	0.1	17859576452.0	1	837102
					2	797903
					3	759787
					4	795577
			0.2	17458182779.0	1	877044
					2	734794
					3	847713
					4	730818
			0.5	15560188811.0	1	995555
					2	601417
					3	602958
					4	990439
ID	44	2	0.1	72402074741.0	1	901753
					2	937353
			0.2	72402074741.0	1	901753
					2	937353
			0.5	62766592211.0	1	690004
					2	1149102
KS	105	4	0.1	23206628505.0	1	697844
					2	700123
					3	769650
					4	770263
			0.2	22280123721.0	1	782388
					2	661105
					3	691499
					4	802888
			0.5	19246870562.0	1	551391
					2	744415
					3	911488
					4	730586
RI	5	2	0.1	infeasible		
			0.2	infeasible		
			0.5	238823827.0	1	660741
					2	436638

Table 5.2: WA redistricting results, with exact solution, 390% threshold

State	n	k	Threshold	Moment of Inertia	District	Population
WA	39	10	3.9	8040514078.0	1	658034
					2	374158
					3	270417
					4	2269675
					5	914814
					6	415342
					7	323800
					8	1003279
					9	845263
					10	630499

The results of the experiments conducted using the model that searches for the exact solution are shown in Table 5.1. Included in the table are the number of counties n, the number of districts k, the objective moment of inertia, and the total population of each of the districts calculated by the model.

From the table it is immediately observable that the redistricting of WA is not possible with any of the chosen thresholds, as we expected from the observations made previously. Other experiments we conducted, reported in Table 5.2, have in fact demonstrated that a threshold of approximately 390% is necessary to make this instance of the problem solvable.

It is also observed that the RI case is solvable only when the threshold is 50%. Hence, in this case, increasing the threshold value has rendered an initially infeasible model operation feasible. The cause of this result can be easily understood by looking at the population of the 5 counties that make up the state. The most populous county, in fact, has 660,741 inhabitants, compared to a total state population of 1,097,379. A minimum threshold of approximately 45% is therefore necessary for the problem to be solvable.

In all other cases, in which the problem has been solved for each threshold used, it can be seen how the value of the objective function decreases as the threshold decreases. This outcome was expected since a higher threshold value reduces the strictness for the population balance among the districts, thus allowing better solutions to be accepted. However, it's important to recognize that this also leads to greater disparities in population sizes between districts.

Finally, noteworthy is the case of AK, where, since $k = 1$, only one solution is possible to solve the problem, regardless of the threshold.

5.2 | Heuristics Solution

Moment of inertia for the heuristic method has been defined based on the two scores:

- Population: the average of the squared deviations of each district's population from the ideal, weighted by a penalty that is heavier if the population is outside the acceptable range. If the population is within the acceptable range, the penalty is lighter.
- Compactness: the inverted sum of all pairwise distances within each district, ensuring that smaller values represent more compact districts. The inversion is achieved by subtracting the actual sum of distances from a maximum possible score, which is calculated based on the greatest distance present in the distance matrix and the total number of possible distances.

With the final moment of inertia being calculated as follows: $I = w_{\text{pop}} \times S_{\text{pop}} + w_{\text{compact}} \times S_{\text{compact}}$, where $w_{\text{pop}} = 2$ and $w_{\text{compact}} = 1$. By doing that we have tried to make the method more focused on the population constraint, however as can be seen in the result table it did not work. The results of this method are hardly dependent on the initial solution for which the only constraint was to make sure everything is connected within the district. Because of the fact that for the new solution to be accepted it does not only have to be better than the previous one or it can be accepted with a probability of $e^{\frac{neigbor\_objective<current\_objective}{T}}$ but also the contiguity of the district cannot be broken, the final populations of district do not differ that much within different threshold. With using *random.seed(0)* for generating

initial solutions for each district, the final solution did not meet the required population constraints for MS: 0.1, 0.2, WA: 0.1, 0.2, 0.5 and RI: 0.1, 0.2. The possible improvement for that issue would be allowing the algorithm to swap multiple counties between the districts at once, so that the population changes between districts are bigger but the contiguity is still fulfilled.

Table 5.3: Results for Heuristics Solutions 1/2

State	n	k	Threshold	Moment of Inertia	District	Population
MS	82	4	0.1*	75978144368.24	1	812979
					2	403130
					3	931156
					4	814014
			0.2*	75982972673.64	1	819857
					2	403130
					3	931156
					4	807136
			0.5	8461609507.83	1	820264
					2	402386
					3	931156
					4	807473
AK	177	1	0.1	156685233.0	0	733391
			0.2	156685233.0	0	733391
			0.5	156685233.0	0	733391
WA	39	10	0.1*	174329291317.78	1	947698
					2	274702
					3	965191
					4	644786
					5	965114
					6	986920
					7	946299
					8	156701
					9	924179
					10	893691
			0.2*	164509287530.09	1	947698
					2	274702
					3	965191
					4	644187
					5	965114
					6	986920
					7	946299
					8	156701
					9	924179
					10	894290
			0.5*	131228616941.78	1	949134
					2	274702
					3	965191
					4	644786
					5	965114
					6	986920
					7	946299
					8	156701
					9	922743
					10	893691

Table 5.4: Results for Heuristics Solutions 2/2

State	n	k	Threshold	Moment of Inertia	District	Population
IA	99	4	0.1	1016545571.73	1	783892
					2	873999
					3	719621
					4	812857
			0.2	1016545571.73	1	783892
					2	873999
					3	719621
					4	812857
			0.5	1003175654.33	1	789646
					2	873999
					3	719621
					4	807103
ID	44	2	0.1	159132757.2	1	926969
					2	912137
			0.2	159132757.2	1	926969
					2	912137
			0.5	159132757.2	1	926969
					2	912137
KS	105	4	0.1	572975955.5	1	753110
					2	730127
					3	726508
					4	683119
			0.2	573858378.3	1	751996
					2	736015
					3	721418
					4	683435
			0.5	580578926.7	1	753110
					2	730127
					3	728195
					4	681432
RI	5	2	0.1*	117022029041.5	1	306799
					2	790580
			0.2*	117022029041.5	1	306799
					2	790580
			0.5	117022029041.5	1	306799
					2	790580

*Threshold values with an asterisk indicate that the final solution did not meet the population constraints.

6 | Group Contributions

Student	Work
Luca Mainardi	Mathematical formulation, exact solution coding, report part regarding results with exact solution, literature study
Laurynas Jagutis	Literature study, solution method, running some experiments for the exact solution, some help in the results section for the exact solution
Ian van de Wetering	Literature study, mathematical formulation, plot generation, and looking into heuristic solutions
Błażej Nowak	Heuristics coding, report part regarding that (Results + help with solution method), literature study
Thomas Warnier	Introduction, literature study and comparisons, looking into exact solutions

7 | References

[1] S. W. Hess, J. B. Weaver, H. J. Siegfeldt, J. N. Whelan, and P. A. Zitlau. Nonpartisan political redistricting by computer. *Operations Research*, 13(6):998–1006, 1965.

[2] R. S. Garfinkel and G. L. Nemhauser. Optimal political districting by implicit enumeration techniques. *Management Science*, 16(8):B–495–B–508, 1970.

[3] Anuj Mehrotra, Ellis L. Johnson, and George L. Nemhauser. An optimization based heuristic for political districting. *Management Science*, 44(8):1100–1114, 1998.

[4] Takeshi Shirabe. Districting modeling with exact contiguity constraints. *Environment and Planning B: Planning and Design*, 36(6):1053–1066, 2009.

[5] Hamidreza Validi, Austin Buchanan, and Eugene Lykhovyd. Imposing contiguity constraints in political districting models. *Operations Research*, 70(2):867–892, 2022.

[6] Micah Altman. Is automation the answer: the computational complexity of automated redistricting. 1997.

[7] William Vickrey. On the prevention of gerrymandering. *Political Science Quarterly*, 76(1):105–110, 1961.

[8] Douglas M. King, Sheldon H. Jacobson, Edward C. Sewell, and Wendy K. Tam Cho. Geo-graphs: An efficient model for enforcing contiguity and hole constraints in planar graph partitioning. *Operations Research*, 60(5):1213–1228, 2012.

[9] Douglas King, Sheldon Jacobson, and Edward Sewell. Efficient geo-graph contiguity and hole algorithms for geographic zoning and dynamic plane graph partitioning. *Mathematical Programming*, 149:425–457, 02 2014.

[10] D. M. King, S. H. Jacobson, and E. C. Sewell. The geo-graph in practice: creating United States Congressional Districts from census blocks. *Computational Optimization and Applications*, 69(1):25–49, January 2018.

[11] Burcin Bozkaya, Erhan Erkut, and Gilbert Laporte. A tabu search heuristic and adaptive memory procedure for political districting. *European Journal of Operational Research*, 144:12–26, 01 2003.

[12] Federica Ricca and Bruno Simeone. Local search algorithms for political districting. *European Journal of Operational Research*, 189(3):1409–1426, 2008.

[13] Michael McDonald and Micah Altman. Bard: Better automated redistricting. *Journal of Statistical Software*, 42, 06 2011.

[14] Wendy K. Tam Cho and Yan Y. Liu. Sampling from complicated and unknown distributions: Monte carlo and markov chain monte carlo methods for redistricting. *Physica A: Statistical Mechanics and its Applications*, 506:170–178, 2018.

[15] Kosuke Imai Benjamin Fifield, Michael Higgins and Alexander Tarr. Automated redistricting simulation using markov chain monte carlo. *Journal of Computational and Graphical Statistics*, 29(4):715–728, 2020.

[16] Paul K Bergey, Cliff T Ragsdale, and Mangesh Hoskote. A simulated annealing genetic algorithm for the electrical power districting problem. *Annals of Operations Research*, 121:33–55, 2003.

[17] Miguel Ángel Gutiérrez-Andrade, Eric Alfredo Rincón-García, Sergio Gerardo de-los Cobos-Silva, Pedro Lara-Velázquez, Roman Anselmo Mora-Gutiérrez, and Antonin Ponsich. Simulated Annealing and Artificial Bee Colony for the Redistricting Process in Mexico. *Interfaces*, 49(3):189–200, May 2019.

[18] Rogelio Z. Rios-Mercado. *Optimal Districting and Territory Design*, volume 284 of *International Series in Operations Research and Management Science*, pages 105–128. Springer, 2020.

[19] National conference of state legislature: Redistricting criteria, 2019.

[20] Federica Ricca, Andrea Scozzari, and Bruno Simeone. Political districting: From classical models to recent approaches. *Annals of Operations Research*, 204:271–299, 04 2013.

[21] Johannes Oehrlein and Jan-Henrik Haunert. A cutting-plane method for contiguity-constrained spatial aggregation. *Journal of Spatial Information Science*, (15):89–120, 2017.

[22] Micah Altman and Michael McDonald. Redistricting by formula: An ohio reform experiment. *American Politics Research*, 46(1):103–131, 2018.