Optimization Meets Participation: Iterative Zone Generation for School Assignment

MOBIN Y. JELOUDAR, IRENE LO, AND KATHERINE L. MENTZER

In U.S. public schools, geographic boundaries significantly influence assignments of students to schools. For example, in 2020 the San Francisco Board of Education approved a policy using multi-school zones with controlled choice to simultaneously improve school diversity and students' proximity to assigned schools. Designing zones for such policies is a computationally and socially complex problem. Algorithmic approaches are needed to provide solutions at scale that balance multiple stakeholder objectives. However, stakeholders often struggle to provide their preferences upfront before seeing feasible zones, limiting their ability to influence algorithmic zone design.

To address these challenges, we propose a stakeholder-in-the-loop framework for preference elicitation and zone design, iterating between using optimization to generate zones, and participatory feedback from stakeholders updating their preferences in reaction to zones. To facilitate real-time iteration between optimization and stakeholder participation, we develop faster computational approaches for multi-school zoning using both math programming and sampling-based methods, and provide interactive participatory tools. Our methods generate preferences for individual stakeholders that can be aggregated using standard social choice approaches, as well as zones with significantly better zone-level statistics than existing benchmarks. Our iterative framework is supporting stakeholder participation in selecting final zones in San Francisco, and is generalizable to other redistricting problems.

1 INTRODUCTION

For U.S. public schools, geographic boundaries have long played an important role in determining which students attend which schools. School districts have used boundaries in student assignment policies to balance competing priorities such as proximity, predictability, and parental choice. One common approach is to use neighborhood boundaries (i.e., 'attendance zones') to determine school assignments, while another promising strategy is controlled choice within multi-school zones [Allman et al., 2022]. However, these approaches can have unintended consequences, such as exacerbating segregation [Holme, 2002, Monarrez, 2023, Owens et al., 2016, Roda and Wells, 2013]. Careful zone design is essential to ensure zone-based policies have their intended effects.

Designing zones for student assignment is a difficult social policy problem that requires overcoming both computational and participatory challenges. Boundary design for neighborhood schools and political districts is NP-hard [Caro et al., 2004, Garfinkel and Nemhauser, 1970]. Designing multi-school zone policies is particularly complex, as zones influence outcomes indirectly through family choice. Computational challenges amplify participatory challenges. Trade-offs are inevitable, since it is typically infeasible to meet all stakeholders' goals. However, stakeholders often cannot articulate desired trade-offs upfront, instead learning preferences by engaging with potential solutions. Presenting feasible maps helps them explore the solution space and provide sufficiently rich preferences for later aggregation. Generating such maps is computationally expensive.

To address the participatory challenges in multi-school zone design, we provide a stakeholder-in-the-loop framework that iterates between: (1) asking stakeholders to update goals and constraints based on generated zones ('Constraint Exploration'); and (2) generating potential zones based on the stakeholder's updated constraints ('Zone Generation'). Through iteration, stakeholders learn their goals while identifying zones that perform well for these goals. This mirrors prior school districting research, which typically provides human-in-the-loop decision-support [Holloway et al., 1975, Ozel et al., 2025], and connects to interactive optimization, where decision-makers iteratively refine feasible solutions and learn about trade-offs through the optimization process [Miettinen et al., 2016]. Our design is also motivated by the participatory design principle that stakeholders better articulate preferences when presented with prototypes and concrete solutions rather abstract prompts [Kensing and Blomberg, 1998, Muller and Kuhn, 1993].

Our framework results in an ensemble of 'good' zone maps illustrating the frontier of trade-offs between quantifiable goals. This consideration set allows stakeholder to specify quantitative preferences that can be aggregated across stakeholders using standard approaches from social choice. It also supports further zone evaluation and discussion about unquantifiable goals and trade-offs. We further support stakeholder participation by developing interactive tools, such as dashboards, interactive spreadsheets, and an LLM-based Constraint Exploration tool.

To address the computational challenges in multi-school zone design, we provide computational methods tailored to the problem. Motivated by neighborhood school zoning, we employ methods based on mathematical programming, accelerated using a novel multilevel optimization approach with interleaved pruning, and a novel recursive optimization approach with overlapping subproblems. Motivated by political redistricting, we tailor Markov Chain Monte Carlo (MCMC) methods to the multi-school zone design problem. We find that our computational methods can rapidly find zone boundaries satisfying given constraints. This enables real-time iteration between zone generation and stakeholder participation, allowing stakeholders to explore zones that satisfy multiple different sets of constraints. As a result, our framework helps the stakeholder generate, in a single session, an ensemble of zone maps that perform well for a wide range of explored goals.

We evaluate our computational methods by comparing the zone maps they produce. When generating larger zones, the multilevel math programming approach consistently produces zone maps with the best zone-level statistics, such as demographic diversity in zones and the balance of students and seats in zones. For scenarios requiring very small zones (e.g., single-school zones), recursive-based solutions yield better results. In contrast, MCMC approaches, such as the Relaxed ReCom algorithm from political redistricting [Charikar et al., 2022], perform well when the number of constraints is limited. However, they struggle to optimize across multiple criteria simultaneously.

We then apply our framework to evaluate the impact of different zone-based policies on student outcomes. We worked with stakeholders in the San Francisco Unified School District (SFUSD) to design zones for each policy considered, then simulated their performance using historical student preferences and school priorities. Comparing the resulting assignments to the current policy, we find that larger zones generally perform better on diversity metrics. Only large zones (maps with 6 zones) reduce both the number of higher-poverty schools and the percentage of historically under-served students in those schools. In contrast, smaller zones (maps with 18 zones) do not improve segregation levels over the current policy, but tend to perform better on proximity measures.

Our framework, simulation, and engagement tools supported the approval of a zone-based policy for SFUSD in 2020 [Allman et al., 2022] and are now guiding stakeholder participation in designing the final boundaries for implementation. Our approach can also be applied to facilitate stakeholder participation in zone design in other school districts or political districting problems.

2 THE MULTI-SCHOOL ZONE DESIGN PROBLEM

In 2018, the San Francisco Unified School District (SFUSD) began redesigning its elementary student assignment policy. The redesign aimed to create a simple, predictable system achieving three main goals. The first goal was *'diversity'*: creating integrated schools reflecting the district's demographically diverse population. The second was *'proximity'*: reducing cross-city travel and enabling more families to be assigned to a school close to home. The third goal was *'predictability'*: reducing families' uncertainty about their child's school assignment. However, the geography of residential segregation in San Francisco makes it challenging to balance diversity and proximity. (By contrast, Berkeley Unified School District also uses a zone-based assignment policy, and the geography of segregation enabled racially diverse and geographically compact hand-drawn zones.)

In December 2020, the San Francisco Board of Education approved a new multi-school zone policy for student assignment. Originally scheduled for 2024–25, implementation is delayed beyond 2026–27 due to district circumstances. The policy divides the district into geographic zones that contain multiple schools, and families are restricted to apply only to General Education programs in their zone. All other programs (e.g. language, special education, K-8) remain citywide. The approved policy did not specify zone boundaries, leaving the design of final zone boundaries to the policy implementation phase. In doing so, the policy effectively deferred the difficult decisions and debates about policy trade-offs that come with proposing specific zone boundaries.

Given the complexity of designing zone boundaries that balance multiple district goals, the implementation phase requires a thoughtful and participatory approach. Early versions of the zone-design tools in this paper were instrumental in the design and approval of the 2020 policy, as they were used to demonstrate that the policy could address all three district goals [Allman et al., 2022]. Our zone design process is also supporting stakeholder participation in designing final zones.

2.1 Challenges in Multi-School Zone Design

Designing multi-school attendance zones presents both computational and participatory challenges. Multi-school zone design is computationally more difficult than the NP-hard problem of designing single-school neighborhood zones. For example, while neighborhood composition directly translates to school composition in neighborhood zoning, for multi-school zones we must also consider the effects of choice within the zone. While standard choice algorithms (e.g., deferred acceptance) can be combined with zones and implemented as integer programs [Roth et al., 1993, Shi, 2022] this leads to extremely large math programming problems ($O(n^4)$ variables, n students) that are intractable to solve directly even with computational tricks and approximations (Appendix A.1).

Multi-school zone design also presents challenges distinct from those in political redistricting. School zones must simultaneously satisfy multiple balance constraints, such as across multiple demographic groups, significantly reducing the number of feasible solutions. The distribution of school seats is highly skewed, with schools concentrated in just 0.7% of census blocks, unlike the relatively continuous distribution of voters in political redistricting. This makes existing redistricting algorithms less effective; for example, local search methods commonly used in redistricting can fail because moving a school across zones can create large imbalances that are difficult to correct with incremental adjustments. In addition, the two problems have different goals: political redistricting seeks to provide a representative set of zone maps, while multi-school zoning seeks to provide an ensemble of 'good' zone maps that illustrate the feasible frontier between competing goals.

The multi-school zone design problem also creates participatory challenges. It is one of many applied optimization problems where a participatory approach is beneficial (and, arguably, necessary) due to the social and computational complexity of the problem. For example, prior work on school districting in operations takes as given the need for human-in-the-loop decision-support systems [Holloway et al., 1975, Overney et al., 2025, Ozel et al., 2025]. Similarly, participatory design emphasizes the need for prototypes, concrete scenarios, and iterative development of algorithmic solutions [Bødker et al., 2004, Kensing and Blomberg, 1998, Muller and Kuhn, 1993]. In these literatures, stakeholder participation in model design is crucial for developing socially responsive models and preventing unintentional policy impacts [Levy et al., 2021], and combining both data-driven and community-driven considerations is often more appropriate than a fully algorithmic approach [Chen et al., 2022, Davis et al., 2022, Delarue et al., 2023].

However, it is difficult to ensure meaningful stakeholder participation in algorithmic multi-school zone design. The typical optimization approach of quantifying desired objectives and constraints upfront is poorly aligned with how stakeholders form and provide preferences. SFUSD stakeholders struggled to state all important constraints upfront due to the many context-dependent constraints (e.g. geographical constraints such as highways/hills). Binding constraints and trade-offs between competing goals often became clear only after feasible solutions were computed and visualized.

The participatory challenges are exacerbated by computational challenges. There are often no feasible solutions that simultaneously achieve multiple goals (e.g., balance constraints across multiple races, short travel distances in all zones). Consequently, the design problem is less about finding an optimal zone map for an individual or group, and more about helping stakeholders explore feasible solutions, articulate preferences among trade-offs, and understand impacts on other stakeholders. In this sense, the multi-school zone design problem exhibits many of the traits that make 'wicked problems' in social policy challenging to address: there are complex social interdependencies; stakeholders have ill-defined yet competing objectives; and it is difficult to define the problem before seeing potential solutions [Churchman, 1967]. These traits necessitate taking a participatory approach to zone design, and inextricably link the participatory and computational challenges.

2.2 Framework

To address both participatory and computational challenges, we propose an iterative stakeholder-in-the-loop framework (Figure 1). Stakeholders identify and refine their preferences by iteratively using generated zones to update desired constraints ('Constraint Exploration'), and generating

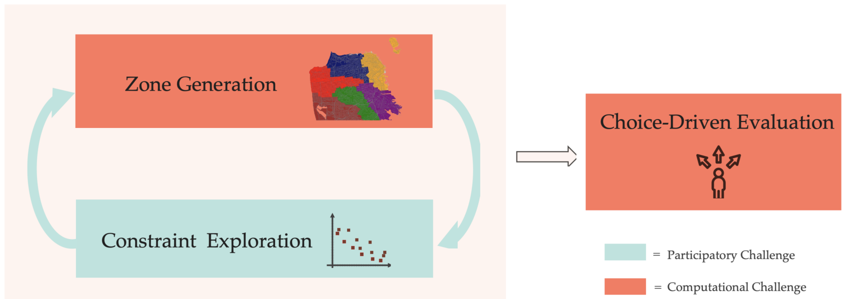


Fig. 1. Iterative stakeholder-in-the-loop framework for multi-school zone design. A Constraint Exploration process is used to help stakeholders learn, refine, and update their desired constraints on zones. A Zone Generation process generates zones that satisfy updated constraints. Stakeholders iterate between these processes to generate a set of feasible zone maps that capture their desired trade-offs between constraints.

new zones using updated constraints ('Zone Generation'). The iterative process addresses the interlinked participatory and computational problems: it simultaneously helps stakeholders explore their desired trade-offs between policy goals, and generates zones that make these trade-offs feasible. This iterative process generates a set of zones that perform well on zone-level metrics. The zones are then further evaluated with school-level metrics that incorporate choice ('Choice-Driven Evaluation'), and the results used to specify quantitative preferences for individual stakeholders, as well as to support stakeholder discussions about achievable policy goals and trade-offs.

In Section 3, we detail the computational methods employed in Zone Generation to create school zones that satisfy stakeholder constraints. We present novel multilevel and recursive math programming approaches, and tailor existing sampling-based methods. In Section 4, we describe how we address the participatory challenges in Constraint Exploration by integrating principles from participatory design with insights on preference formation from the consumer search literature.

The 'Choice-Driven Evaluation' component of our framework allows us to tackle the computational challenge of zone design with choice. We reduce the computational complexity in the zone-generation and constraint-exploration stage by focusing on zones that perform well with respect to zone-level measures of district goals. In Choice-Driven Evaluation, we then simulate the choices of students within each zone by running deferred acceptance with diversity reserves, district priorities and simulated preferences [Gale and Shapley, 1962]. We use this to evaluate how well the resulting assignment meets quantifiable measures of policy goals. Finally, we elicit quantitative preferences for zone maps from stakeholders that can be aggregated using standard social choice approaches.

3 COMPUTATIONAL METHODS FOR SCHOOL ZONE GENERATION

In this section, we explore computational approaches for generating school zones. Our goal is to generate an ensemble of solutions that represent different value judgments while satisfying constraints on balancing demographics, students, and schools. In Section 3.1, we adapt math programming approaches commonly used for neighborhood school zoning and political redistricting to multi-school zoning. In Section 3.2, we explore sampling-based methods from political redistricting.

Preliminaries. All balance constraints in this section were identified via our Constraint Exploration process (see Section 4 for details). All computational methods in this section create zones by combining small geographic units into zones. In districts like SFUSD, these units can be census blocks, census block groups, or single-school neighborhood zones (known in SFUSD as 'attendance areas'). San Francisco has roughly 6,000 census blocks, grouped into 579 block groups (each containing about 8–12 blocks), and 59 attendance areas. Boundaries between block groups and attendance areas do not always align.

3.1 Mathematical Programming Approaches

Mathematical programming approaches are commonly used in both neighborhood school zoning [Caro et al., 2004, Gillani et al., 2023] and political redistricting [Garfinkel and Nemhauser, 1970, Gurnee and Shmoys, 2021, Mehrotra et al., 1998], and typically employ integer programming to assign geographic units to zones while meeting district-defined objectives and constraints. For SFUSD, directly solving the multi-school zones math program for the entire school district is prohibitively computationally expensive. To address this, we introduce two approaches, adapted from existing methods to reduce problem size and computational burden. The first approach is a multilevel method that utilizes pruned solutions from less granular levels (e.g., attendance area) to initialize partial solutions at more granular levels (e.g., block). The second approach employs recursive methods with overlapping subproblems to divide the district into many small zones.

Multilevel algorithms are a class of graph-based methods that iteratively coarsen a graph (by merging adjacent nodes), solve a problem on the smaller graph, and then project the solution back to the original graph [Barnard and Simon, 1994, Blondel et al., 2008, Hendrickson et al., 1995, Karypis and Kumar, 1998]. They are highly effective for graph partitioning problems in applications such as VLSI design and image processing [Buluç et al., 2016]. The multilevel paradigm has been applied to political redistricting [Swamy et al., 2023], the traveling salesman problem [Walshaw, 2002], and graph drawing [Walshaw, 2001]. We adapt the multilevel approach to school zoning and demonstrate that solving a less granular problem can provide a valuable starting point for more refined solutions.

Recursive algorithms have been employed to increase computation speed in political redistricting. [Gurnee and Shmoys, 2021] propose a two-stage optimization approach using a randomized divide-and-conquer column-generation heuristic. They generate many zones by exploiting the compositional structure of graph partitioning problems and then optimize over this set. Other recursive approaches include shortest split-line [Kalcsics et al., 2005], diminishing halves [Spann et al., 2007], and other divide-and-conquer approaches [Levin and Friedler, 2019]. We introduce a recursive approach to school zoning where subproblems are allowed to *overlap*, and show that it provides a viable approach, particularly when zones are built from fewer geographic units.

3.1.1 Mixed Integer Programming (MIP) Formulation. We formulate the school zoning problem as an integer program. Existing neighborhood zoning approaches use integer programming to connect geographic units to zone 'centroids', which are pre-defined geographical units in each zone that serve as a reference point for ensuring contiguity [Caro et al., 2004, Gillani et al., 2023]. Each zone has a natural centroid: the school. We provide a formulation of multi-school zoning where each zone contains multiple schools. Our formulation allows centroids to be selected either exogenously, or endogenously as part of the optimization (Appendix A.3). Our formulation can also be modified to jointly optimize zone boundaries and program placement in schools (Appendix A.7).

Let U denote the set of geographic units and Z the set of zones, each associated with a centroid unit $z \in U$. The units can be attendance areas, block groups, or blocks. We had the computational capacity to solve the MIP for ≈5 zones and ≈800 units, which is sufficient for attendance area units; for smaller units we use the multilevel and recursive methods in Sections 3.1.3 and 3.1.2.[1]

[1]Solving the MIP formulation at the attendance area level took ≈1, 6, and 20 minutes for 4-, 7-, and 14-zone problems respectively. At the block group level, solving for 4 zones took ≈4 minutes; solving for more zones at the block group level

For each unit $u \in U$ and zone $z \in Z$, let the decision variable $x_{u,z} \in \{0,1\}$ indicate whether unit u is assigned to zone z. For each unit u, let $N(u)$ be the geographic units that neighbor u (i.e., sharing a border in the 2020 census map). We will consider covariates associated with each unit u, including the number of schools Sch_u, seats $Seat_u$, and students (or population) P_u in the unit. The integer program aims to allocate geographic units to zones with the following objectives and constraints:

Feasibility. The following constraint ensures that every unit belongs to exactly one zone.

$$\sum_{z \in Z} x_{u,z} = 1 \qquad \forall u \in U. \tag{3.1}$$

Contiguity. Contiguity ensures that every part of a school zone is reachable from any other part without leaving the zone. Contiguity is important for school proximity and zone interpretability, and was specified by the San Francisco Board of Education as a necessary constraint. We tested multiple approaches to contiguity, and adapt a contiguity constraint formulation from [Caro et al., 2004], which requires pre-defined zone centroids for building each zone.[2] Let $d_{u,z}$ be the Euclidean distance between unit u and the centroid of zone z (2020 census data). Our contiguity constraint is:

$$x_{u,z} \leq \sum_{v: v \in N(u), d_{v,z} \leq d_{u,z}} x_{v,z} \qquad \forall u \in U, z \in Z. \tag{3.2}$$

Equation (3.2) ensures that unit u is assigned to zone z only if there is a 'path' of closer neighboring units connecting u to the zone centroid. While this constraint is sufficient for contiguity, it is not necessary, and may exclude valid contiguous zone maps; we thus enforce it only for units u and zones z where u has at least one neighbor closer to z (see Appendix A.4 for details).

Compactness. A common preference shared by district decision-makers, staff, and families in school redistricting is for *compact* zones. Compactness requires zones to have a reasonable shape, avoiding 'snake-like' paths. Several measures have been proposed to quantify zone compactness, such as total zone perimeter, comparing each zone's area and perimeter [Duchin and Tenner, 2024, Young, 1988], summing the distance from each geographic unit to the center or centroid of its zone, or counting the number of cut edges in the graph representation (i.e., edges whose endpoints are in different zones) [Dobbs et al., 2023, Validi and Buchanan, 2022]. We encouraged compactness by minimizing the number of cut edges as in [Dobbs et al., 2023], as follows:[3]

$$Obj_{IP} \;=\; \min \sum_{u \in U,\; v \in N(u)} b_{u,v} \tag{3.3}$$

$$b_{u,v} \geq |x_{u,z} - x_{v,z}| \qquad \forall u \in U, v \in N(u), z \in Z. \tag{3.4}$$

The *boundary cost* $b_{u,v} \in \{0,1\}$ indicates whether units u and v are assigned to different zones. The objective (3.3) minimizes the total number of neighboring units assigned to different zones. Constraint (3.4) ensures that $b_{u,v}$ is 1 if and only if units u and v are assigned to different zones.

or any zones at the block level was not computationally feasible. The multilevel approach significantly sped up computation, taking about 11 minutes per iteration for 7-zone maps and 35 minutes for 14-zone maps.

[2]Alternative approaches to contiguity include using different distance measures, constraints on the dual graph [King et al., 2012], cutting plane methods [Oehrlein and Haunert, 2017, Salazar-Aguilar et al., 2011], and solving without full contiguity constraints and modifying discontiguous solutions [Gentry et al., 2015]. We found the formulation in [Caro et al., 2004] sufficient for visual performance and computational efficiency. The cutting plane approaches and approach in [Gentry et al., 2015] were too computationally slow and too manual respectively. We also tested alternative distance measures in the contiguity formulation, including driving distance and graph distance, but found these significantly reduced the number of contiguous zones.

[3]We tried different compactness measures, including combinations of cut edges and squared distance from centroids. In our setting, using cut edges alone yielded contiguous zones that appeared the most visually compact to stakeholders.

Balance Constraints. Another set of constraints common in school redistricting are balance constraints. To support planning and logistics, school zones should be of a similar size, and provide sufficient seats for students in the zone. In addition, when working with SFUSD stakeholders, we found that stakeholders were much more able to understand and articulate trade-offs between policy goals when they were articulated as constraints with thresholds that could be loosened/tightened, rather than objectives with weights that could be tuned. We therefore encode many of the district's goals as balance constraints. For example, to promote diversity and equity, we include constraints enforcing demographic similarity between zones.[4] The balance constraints are as follows.

$$\left|\sum_{u\in U}\frac{\mathrm{Sch}_u}{|Z|}-\sum_{u\in U}\mathrm{Sch}_u.x_{u,z}\right|\leq 1 \qquad \forall z\in Z, \tag{3.5}$$

$$\left|\sum_{u\in U}(\mathrm{P}_u-\mathrm{Seat}_u)x_{u,z}\right|\leq \alpha_{capacity}\sum_{u\in U}\mathrm{P}_u x_{u,z} \qquad \forall z\in Z, \tag{3.6}$$

$$\left(\frac{F}{N}-\alpha_{SES}\right)\cdot\sum_{u\in U}\mathrm{P}_u x_{u,z}\leq\sum_{u\in U}f_u x_{u,z}\leq\left(\frac{F}{N}+\alpha_{SES}\right)\cdot\sum_{u\in U}\mathrm{P}_u x_{u,z} \qquad \forall z\in Z, \tag{3.7}$$

$$\left(\frac{R^k}{N}-\alpha_k\right)\cdot\sum_{u\in U}\mathrm{P}_u x_{u,z}\leq\sum_{u\in U}R^k_u x_{u,z}\leq\left(\frac{R^k}{N}+\alpha_k\right)\cdot\sum_{u\in U}\mathrm{P}_u x_{u,z} \qquad \forall z\in Z,k\in K. \tag{3.8}$$

Recall that Sch_u, P_u and Seat_u are the number of schools, students (population), and seats in unit u. Constraint (3.5) ensures all zones have roughly the same number of schools (±1); which SFUSD decision-makers prioritized over exact student balance.[5] Constraint (3.6) limits the difference between students and seats in each zone so that shortages and overages do not exceed $\alpha_{capacity}$ of its total student population. Because zones are set for multiple years and student populations vary yearly, eliminating all shortage and overage is infeasible. In SFUSD, we set $\alpha_{capacity}$ between 15% and 30% based on zone size (e.g.,15% for 10-school zones and 30% for 3-school zones).

Constraints (3.7) and (3.8) capture demographic balance across schools. Specifically, for each $u\in U$, the quantity f_u is the total socio-economic score of students in unit u, and $F := \sum_{u\in U} f_u$ is the total score over all units. In line with the education policy literature [see, e.g. Domina et al., 2018, Harwell and LeBeau, 2010, Nicholson et al., 2014], our socioeconomic score is given by the percentage of students in the unit eligible for Free or Reduced Priced Lunch (FRL), a common proxy for student disadvantage in U.S. public schools. Similarly, for a set of races $K :=$ {White, Asian, Latinx},[6] and for each unit u and race $k\in K$, the quantity R^k_u is the number of students of race k in unit u, and $R^k := \sum_{u\in U} R^k_u$ is the total number of students of race k. Finally, $N := \sum_{u\in U}\mathrm{P}_u$ is the total number of students in all units. Constraint (3.7) ensures the average socioeconomic score of students in each zone is within α_{SES} of the average socioeconomic score of students in the whole district, and Constraint (3.8) ensures the proportion of students of a given race k in each zone is within α_k of the proportion in the district. For SFUSD, we set α_{SES} and α_k to be a value between 12% and 15%, depending on zone size: for larger zones, we use stricter balance constraints as it is easier to achieve socio-economic and racial balance; for smaller zones we gradually relax constraints.

[4] While zone composition is not the final target measure of diversity, as choice allows the composition of zones and schools to differ, district stakeholders still believed it was important to aim for some level of demographic similarity between zones.
[5] Unlike in political redistricting, where '1 person 1 vote' is constitutionally mandated, in school zoning students counts don't need to be exactly balanced across zones.
[6] White, Asian, and Latinx are the largest racial groups in SFUSD. Black students comprise only ≈ 5% of the student population, making it challenging (and potentially undesirable) to balance them across zones.

Centroid Selection. Our primary goal in working with SFUSD was to enable fast computation of many zone options. We thus exogenously selected centroids by considering high-popularity schools dispersed throughout the city as centroids and building zones around them. In collaboration with SFUSD staff, we generated and explored 90 different sets of centroids. We also provide an alternative approach that endogenously selects centroids by treating them as variables (Appendix A.3).

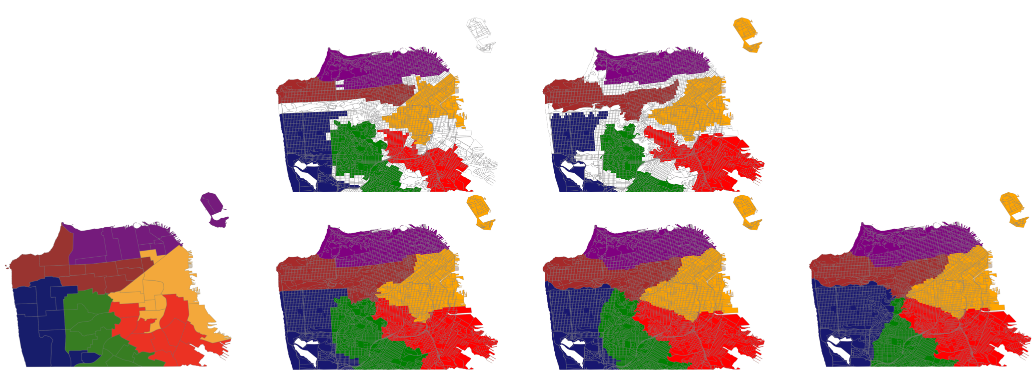
3.1.2 Multilevel Zone Design. We aim to solve the MIP from Section 3.1.1 using census blocks. This results in 23,847 variables and 116,645 constraints, which is computationally intractable. We thus employ a multilevel optimization procedure with interleaved pruning (Algorithm 1, Figure 2), which progressively refines the solution at different granularity levels, trimming the borders of zones in solutions at lower levels of granularity and re-solving at higher granularity. Our levels of granularity are attendance areas, then block groups, then blocks.

ALGORITHM 1: Multilevel Zone Design

Initialization: Solve the MIP from Section 3.1.1 using low-granularity geographic units.
Iterative Optimization: Iteratively use the solution from the previous step, at a lower-granularity level, to find a solution at a higher-granularity level, as follows.

- ***Mapping & Trimming:*** Initialize partial solution using next level of granularity. Map each unit to zone of corresponding lower-granularity unit from previous step. Trim zone assignment for any units with few adjacent units in the same zone. Tighten balance constraints.
- ***MIP With Initialized Variables:*** Initialize the MIP variables using the trimmed partial solution. Solve for remaining variables, ensuring constraints from Section 3.1.1 are satisfied.

Local Search: Improve solution at highest-granularity level using local search (Appendix A.5).



(a) Initial Zoning (Attendance area) (b) Iterative Optimization (Block Group) (c) Iterative Optimization (Block) (d) Local Search (Block)

Fig. 2. Steps in multilevel zone design (Algorithm 1). In Figure (a), an initial solution is obtained by solving the MIP for attendance areas. In Figures (b) and (c), a partial solution is initialized at the next level of granularity and the MIP solved for the remaining variables. In Figure (d), local search is applied to smooth the boundaries of zones.

The trimming step promotes zone compactness by removing isolated units, enabling the optimization model to find improved solutions at the more granular level. Units near the zone centroid remain in the same zone, while boundaries for more distant units can vary considerably. For SFUSD, we first trim the units which have over 40% of their neighboring units in different zones, and then trim units which lack a contiguous path to their zone's centroid. Smaller units in later stages provide greater flexibility, enabling tightening of balance constraints by about 8% at each step.

3.1.3 Overlapping Recursive Zone Design. When desired zones are small (e.g. neighborhood zones), low-granularity geographic units cannot be used to initialize multilevel zone design. We thus introduce an overlapping recursive algorithm (Algorithm 2, Figure 3) that can be used instead of recursive zoning to speed up computation. Recursive zoning algorithms typically divide the problem into smaller non-overlapping subproblems, solve them separately, and combine the solutions at the end [Gurnee and Shmoys, 2021, Levin and Friedler, 2019]. They tend to perform poorly in multi-school zoning due to the problem's highly constrained nature (Appendix A.6). Our overlapping recursive algorithm divides the district into overlapping regions, where overlapping areas can be re-zoned. It can thus 'grow' a viable solution from a pruned subset of an initial solution, and is not overly constrained by poor initial divisions.

The integer program formulation for overlapping recursive zoning is similar to the MIP described in Section 3.1.1. The primary difference is in feasibility constraints: overlapping recursive zoning divides the district into overlapping subsections, and enforces extended feasibility constraints on the regions of overlap. Formally, let S_{opt} represent the centroids targeted for zone optimization (Figure 3a). Let S_{ext} denote an expanded set of centroids that includes S_{opt} and additional 'buffer' centroids; the buffer centroids $S_{ext} \setminus S_{opt}$ are strategically selected to create buffer zones around the S_{opt} zones, separating them from the remaining zones in the district (Figure 3b). The overlapping recursive integer programming formulation uses the objective (3.3)-(3.4), contiguity constraint (3.2), and balance constraints (3.5)-(3.8) from the MIP in Section 3.1.1, defined for all $u \in U_{flex}$ and $z \in S_{ext}$. In addition, we modify the feasibility constraint (3.1) as follows.

Extended Feasibility. Let U_{must} be the set of units within a certain distance (e.g., 1 mile) from centroids in S_{opt}. These units must be assigned to zones with centroids in S_{ext} (Figure 3c). Let $U_{flex} \supseteq U_{must}$ be a set of flexible units within a certain distance (e.g., 4 miles) from centroids in S_{opt}, that may or may not be assigned to zones with centroids in S_{ext} (Figure 3d).

$$\sum_{z \in S_{ext}} x_{u,z} = 1 \quad \forall u \in U_{must}, \tag{3.9}$$

$$\sum_{z \in S_{ext}} x_{u,z} \leq 1 \quad \forall u \in U_{flex}, \tag{3.10}$$

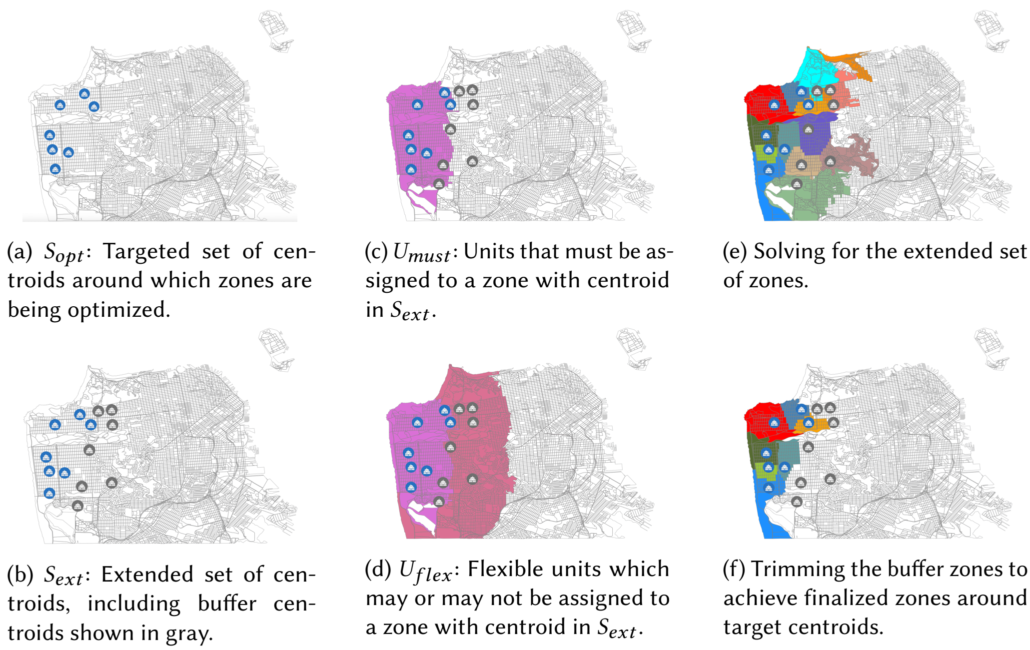
$$\sum_{z \in S_{ext}} x_{u,z} = 0 \quad \forall u \in U \setminus \left(U_{must} \cup U_{flex}\right). \tag{3.11}$$

Constraint (3.9) ensures that each unit in U_{must} is assigned to a zone with centroid in S_{ext}. Constraint (3.10) considers whether assigning U_{flex} units to S_{ext} centroids improves the objective. Constraint (3.11) ensures that units outside of U_{flex} are not assigned to zones with centroids in S_{ext}.

The overlapping recursive algorithm (Algorithm 2) maintains compactness through overlapping subproblems. The centroids S_{ext} and units U_{flex} for the current subproblem overlap with the target centroids and flexible units that will be assigned in a subsequent subproblem. The extended school sets (S_{ext}) serve as buffers, facilitating smooth transitions between adjacent subproblems.

ALGORITHM 2: Overlapping Recursive

- Select S_{ext}, S_{Opt}, U_{must} and U_{flex}
- Solve the MIP in Section 3.1.3 for S_{ext}, S_{Opt}, U_{must} and U_{flex}
- Fix and finalize the zones around centroids in S_{opt}
- Proceed to solve the next subproblem (e.g., moving from West to East).



(a) S_{opt}: Targeted set of centroids around which zones are being optimized.

(c) U_{must}: Units that must be assigned to a zone with centroid in S_{ext}.

(e) Solving for the extended set of zones.

(b) S_{ext}: Extended set of centroids, including buffer centroids shown in gray.

(d) U_{flex}: Flexible units which may or may not be assigned to a zone with centroid in S_{ext}.

(f) Trimming the buffer zones to achieve finalized zones around target centroids.

Fig. 3. Creating zones around target centroids in overlapping recursive zone design

3.2 MCMC Approach

In this section, we adapt *ensemble-based methods* to the multi-school zoning problem. These methods are commonly used in political redistricting to sample ensembles of possible plans [Autry et al., 2021, DeFord et al., 2021], allowing stakeholders to contextualize proposed plans and detect gerrymandering [Chikina et al., 2020, Fifield et al., 2020, Herschlag et al., 2017]. While ensemble-based methods can be used to create school zone maps, they need to be adapted to address the unique challenges of multi-school zoning (cf. Section 2.1). We describe the MCMC methods typically used in political redistricting, and how we tailor them to multi-school zoning.

3.2.1 Sampling Methods for Redistricting. Sampling methods like Metropolis Markov Chain Monte Carlo (MCMC) are typically used in political redistricting to generate large, diverse ensembles of possible zoning plans [Autry et al., 2021, DeFord et al., 2021]. The political redistricting literature focuses on proving these methods achieve *representative* sampling, which is crucial for assessing the fairness of a given map, and characterizing mixing times. Such methods have been successfully employed in legal challenges to partisan gerrymanders at state and federal levels and used as evidence in court cases [DeFord et al., 2021, Herschlag et al., 2020a, Zhao et al., 2022].

A key aspect of applying MCMC methods is the sampling distribution over possible partitions. MCMC methods model redistricting as a graph partition problem and provide random walks to transition between neighboring partitions. Different random walks induce different sampling distributions. One family of random walks used for sampling in redistricting is *Flip Walks*, which, at each step, randomly selects a geographic unit on the boundary of a zone and reassigns it to an adjacent zone while preserving zone contiguity. This naive approach samples from the uniform distribution over all partitions of geographic units into contiguous regions satisfying population balance constraints. Although simple to implement and computationally fast, Flip walks suffer from slow mixing times (i.e. time to sample the space of partitions), and a tendency to generate non-compact zones [DeFord et al., 2021]. It is thus unlikely to be suitable for iterative multi-school zoning.

Recombination (ReCom) is another family of random walks that was introduced by [DeFord et al., 2021] to address the limitations of Flip walks. Recom improves mixing times by making more substantial changes to the partition at each step; we detail these changes in Section 3.2.2. In addition, the probability that ReCom traverses to a neighboring zone map depends on the product of the number of spanning trees within each of its zones. This has the desirable property of concentrating probability mass on more compact zones, as compact zones tend to have a larger number of spanning trees [DeFord and Duchin, 2019, Procaccia and Tucker-Foltz, 2022]. Recent studies show that ReCom can efficiently explore the space of valid plans and effectively redraw electoral district boundaries in various states [Akitaya et al., 2023, Dobbs et al., 2023]. We adapt a variant of the ReCom algorithm and explore its performance on the multi-school zoning problem.

Temperature-based variants of the Flip and ReCom approaches, such as simulated annealing and parallel tempering [Bangia et al., 2017], have also been proposed to address the mixing time challenges encountered by MCMC approaches. However, our simulations suggest the effectiveness of these techniques is limited when applied to the school zoning problem (Appendix A.8). Other MCMC approaches have been proposed to address the limitations of traditional methods including Forest ReCom [Autry et al., 2008, 2021], lifting [Herschlag et al., 2020b, Vucelja, 2016] and Sequential Monte Carlo [McCartan and Imai, 2023]. However, school zoning's multiple balance constraints require tailored sampling approaches, which we explore next.

3.2.2 MCMC Approaches for School Zoning. To study the efficacy of MCMC approaches in school redistricting, we implement Relaxed ReCom (Algorithm 3), a variant of the ReCom algorithm introduced by [Charikar et al., 2022], and tailor a sampling distribution for school zoning.

School zoning and political redistricting can be modeled as a graph partition problem [Schaeffer, 2007]. The problem is represented by a graph $G = (U, E)$, where the vertex set U is the set of geographical units (e.g., census blocks) and E is the set of edges connecting adjacent units. A zoning $(Z_1, ..., Z_k)$ is a partition of U into k zones where each zone is a connected subgraph of G. In this section, for $U' \subseteq U$ we let $G[U']$ denote the subgraph of G induced by U'.

The ReCom walk proceeds by iterating over the following steps. (1) Randomly select and merge two adjacent zones. (2) Form a random spanning tree of the induced subgraph of the merged zones.[7] (3) Cut a uniformly random edge of the spanning tree to create two new zones. (4) If the new zones meet balance constraints, update the zoning plan. Relaxed ReCom modifies this approach by: using a biased distribution when selecting the cut edge in (3); and removing the balance constraint check in (4), allowing for broader exploration. The balance constraints are used at the end to filter the ensemble of generated solutions. While more computationally intensive per step than ReCom or Flip, Relaxed ReCom has the advantage of rapid mixing properties.[8]

Biased Spanning Tree Distribution: To implement Relaxed ReCom for the multi-school zone problem, we define a sampling distribution μ^b over all possible school zonings, and set $\mu = \mu^b$ in Algorithm 3. Given a graph $G = (U, E)$ and number of zones k, the distribution μ^b is defined by

$$\mu^b(Z_1, \ldots, Z_k) = \prod_{i=1}^{k} T(Z_i) \cdot |\mathrm{Sch}_i|^{w_{\mathrm{Sch}}} \cdot |Z_i|^{w_{\mathrm{size}}} \cdot |\mathrm{Shortage}_i|^{w_{\mathrm{Shortage}}} \cdot |\mathrm{FRL}_i|^{w_{\mathrm{FRL}}},$$

where $Z_1 \cup ... \cup Z_k = U$ is a zoning of U, $T(Z_i)$ is the number of spanning trees in $G[Z_i]$, $|\mathrm{Sch}_i|$ and $|Z_i|$ are the number of schools and units in zone i respectively, $\mathrm{Shortage}_i$ is the percentage of

[7] Random spanning trees of graphs can be formed efficiently using Wilson's algorithm [Wilson, 1996].

[8] [Charikar et al., 2022] demonstrates rapid mixing of Relaxed ReCom for grid graphs with balanced homogeneous weights, though extending these guarantees to more general graphs remains an open question.

students in zone i exceeding available seats, FRL_i is the proportion of students in zone i eligible for Free or Reduced Priced Lunch, and $w_{Sch}, w_{size}, w_{Shortage}, w_{FRL} \geq 0$ are predefined weights balancing the distribution of schools, zone size, shortage, and FRL-eligible students across zones.

ALGORITHM 3: Relaxed ReCom [Charikar et al., 2022]

Input: Graph $G = (U, E)$, zoning $(Z_1, ..., Z_k)$, a spanning tree for each zone $(T_1, ..., T_k)$, and a distribution μ over possible zonings of G.
Output: A new zoning $(Z_1^{new}, ..., Z_k^{new})$, and a new spanning tree for each zone $(T_1^{new}, ..., T_k^{new})$.
In each iteration:

(1) Add a uniformly random edge e that merges two adjacent zones Z_i and Z_j.
(2) Form the spanning tree $T_i \cup T_j \cup \{e\}$ of the induced graph $G[Z_i \cup Z_j]$ of the merged zone.
(3) Cut a random edge f of $T_i \cup T_j \cup \{e\}$ with probability proportional to $\mu(Z_1^f, \ldots, Z_k^f)$,[9] where Z_s^f is defined as follows:
   - Let T_i' and T_j' be the trees created by removing f from $T_i \cup T_j \cup \{e\}$.
   - Let Z_i' and Z_j' be the vertex sets of T_i' and T_j' respectively.
   - Let $Z_s^f := \begin{cases} Z_s' & \text{if } s = i \text{ or } j \\ Z_s & \text{otherwise,} \end{cases}$ and let $T_s^f := \begin{cases} T_s' & \text{if } s = i \text{ or } j \\ T_s & \text{otherwise.} \end{cases}$
(4) Set new output $Z_i^{new} := Z_i^f$ and $T_i^{new} := T_i^f$ $\forall i$, corresponding to the selected cut edge f.

If we set all weights w to zero, the biased spanning tree distribution reduces to the spanning tree distribution. The term $\prod_{i=1}^{k} |\text{Sch}_i|$ increases as the number of schools across zones becomes more balanced (and similarly for zone size, shortage, and FRL). Consequently, the biased distribution assigns the highest probability to zone maps where different zones have similar numbers of schools, numbers of students, shortage, and proportion of students eligible for FRL. As w_{size} tends to infinity, the biased spanning tree distribution approximates the balanced spanning tree distribution, effectively imposing a 'soft' constraint on the balance of zone sizes.

4 CONSTRAINT EXPLORATION

In this section, we present how we address the participatory challenges in the Constraint Exploration component of our stakeholder-in-the-loop framework. Given the optimization approach to generating zones, our goal is to allow stakeholders to learn their desired trade-offs between feasible sets of constraints. We focus on learning optimization constraints, rather than weights for different components of an objective, as this was more intuitive and approachable for SFUSD stakeholders.

This framing builds on work in constraint-based preference elicitation, which shows that decision-makers often find it easier to specify acceptable and unacceptable conditions in the form of constraints, rather than assigning precise weights to competing objectives [Boutilier et al., 2005]. It also connects to the literature on interactive multiobjective optimization and decision aiding, where iterative solution generation helps participants clarify their goals and identify which constraints or trade-offs they wish to enforce [Miettinen et al., 2016, Tsoukiàs, 2008]. Such an approach has been used recently in multiple school districts to facilitate community participation in algorithmic policies like districting and scheduling [Delarue et al., 2023, Ozel et al., 2025]. To incorporate the varied goals of different stakeholders, we draw on principles from *Participatory Design* (PD), a framework for creating a hybrid space where diverse stakeholders can meaningfully contribute to designing socio-technical systems [Bødker et al., 2004, Kensing and Blomberg, 1998, Muller and Kuhn, 1993]. Our approach to the Constraint Exploration problem was used iteratively with our

[9] We employ the down-up walk method from Anari et al. [Anari et al., 2021] to sample the new cutting edge f at each step. This method takes as input any distribution over graph partitions and samples edges based on that distribution.

Zone Generation methods to identify feasible constraint sets and zone maps for those constraint sets for SFUSD.

While the framework presented in Section 2.2 suggests a single Constraint Exploration process, our close collaboration with the district revealed the need to engage human decision-makers before every computational step. Thus, there are three main participatory components that support our stakeholder-in-the-loop framework (Figure 4): identifying metrics that align with stakeholders' preferences ('Metric Elicitation'); exploring preferences by finding a set of well-performing zones and feasible constraints ('Consideration Set Formation'); and selecting the set of constraints from the consideration set that best align with the stakeholder's learned objectives ('Constraint Selection'). In the rest of this section, we provide further details about each of these stages.

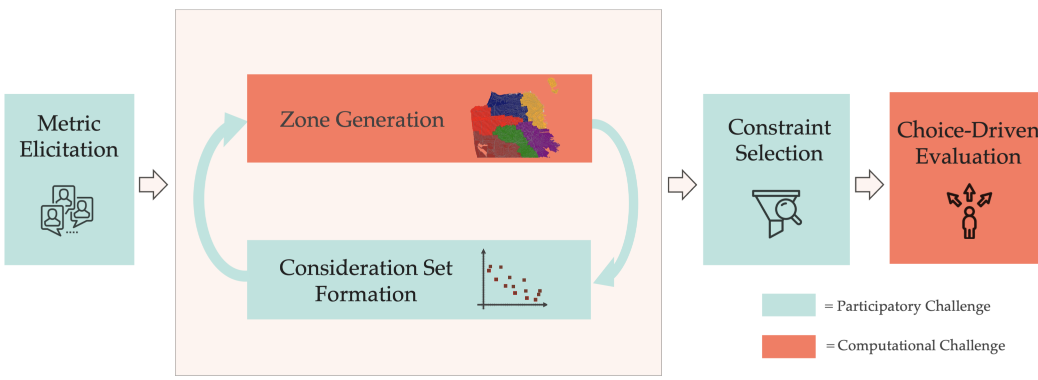


Fig. 4. Participatory components in our stakeholder-in-the-loop framework. Support for stakeholder participation was needed in Metric Elicitation, the iterative loop of Consideration Set Formation, and also forming initial preferences via Constraint Selection.

4.1 Metric Elicitation

The starting point of our participatory process is identifying metrics to be used in Constraint Exploration. These metrics should be simple yet flexible enough to capture the goals, needs, and desires of many different stakeholder groups. At the same time, the metrics need to be quantifiable and usable in Zone Generation algorithms. To identify metrics, SFUSD organized several focus groups, an effective method for involving a broad range of stakeholders in initial stages of participatory design [Muller and Kuhn, 1993]. The groups included participants from various organizations and stakeholder segments.[10] We engaged focus group participants by starting with the metrics and constraint sets used to develop the 2020 Board policy, and a set of well-performing zone maps developed to support that policy. Participants provided feedback on the metrics to be balanced, allowing us to introduce new types of constraints and adjust the format of existing constraints. Selected metrics are described in Section 5.2. Further details on focus groups can be found in Appendix B.1.

4.2 Consideration Set Formation

The primary iterative component of our Constraint Exploration process is Consideration Set Formation, where stakeholders iteratively use generated zones to update desired constraints. This

[10]Individual focus groups were held with the SFUSD Educational Placement Center (EPC), the San Francisco Municipal Transit Agency (SFMTA), the San Francisco Department of Children, Youth, and their Families (DCYF), SFUSD's Superintendent's Executive Cabinet, parent leaders, and former SFUSD Student Assignment staff.

component is motivated by a long tradition in operations research of interactive optimization [Miettinen et al., 2016], and the Participatory Design (PD) approach of using prototypes to enable non-technical stakeholders to provide input on technology design [Bødker et al., 2004, Kensing and Blomberg, 1998, Muller and Kuhn, 1993, Sanders and Stappers, 2014]. We found that stakeholders could more effectively quantify their desired trade-offs when presented with actual zone maps, rather than abstract goals or measures of distance or diversity. The iterative process allowed participants to explore the feasible space and learn their preferred constraints.

In the Consideration Set Formation stage, stakeholders build a Pareto frontier of constraint sets that are achievable via different zone maps. This formation of a well-performing ensemble of zones mirrors ‘consider-then-choose’ behavior from consumer search, where consumers form a ‘consideration set’ before making their final choice [see, e.g., Aouad and Segev, 2021, Hauser and Wernerfelt, 1990]). Each constraint (e.g. student balance, seat balance) consists of a metric, and a threshold allowing a certain amount of deviation from exact balance (e.g., 5% seat shortage). Borrowing from the consumer search literature, we refer to a set of constraints and their respective thresholds as a *position*. Participants explore various positions, tightening some constraints and relaxing others to search through the feasible space. We are interested in feasible positions that are Pareto optimal, i.e., there are no other feasible positions where all constraint thresholds are more restrictive, with at least one strictly more restrictive. We explored two approaches for Consideration Set Formation.

Approach 1: Sequential Process. Stakeholders build a consideration set through sequential screening. We start with a set of constraints yielding a feasible solution and gradually tighten thresholds. When we reach an infeasible position, we have approximately identified a Pareto-improving map and add it to the consideration set. Stakeholders decide whether to stop the screening, finalizing their consideration set, or query a new position, discussing which constraints they wish to loosen. This iterative process reveals the frontier of feasible solutions through multiple parallel search paths, each exploring different constraint combinations. We provide an illustration in Appendix B.2.

Approach 2: Automatic Tool Using Large Language Models. To streamline the constraint exploration process and empower stakeholders to independently investigate the solution space and build their consideration set, we prototyped a generative-AI-powered interactive user interface (UI) (Figure 5). This web-based application allows stakeholders to request, in natural language, additional constraints or desired changes to constraints. The tool then uses large language models to automatically convert requests into optimization constraints [AhmadiTeshnizi et al., 2023], add or change these constraints in the optimization module, generate new zones, and display resulting zoning solutions to the stakeholder in real-time. The tool significantly accelerated the iteration cycle by removing the researcher from the loop [Delarue et al., 2023].

4.3 Constraint Selection

Once the consideration set is formed, stakeholders proceed to Constraint Selection, where they analyze zone maps in their consideration set. This is also the first step towards forming quantitative preferences. To support this process, we developed an interactive dashboard for exploring optimization-generated solutions. The dashboard allows stakeholders to account for additional information such as accessibility to public transportation and other non-quantitative social objectives, and to choose the zones and corresponding sets of constraints that best aligns with their objectives.

The dashboard consists of three selection stages, each targeting one of the district’s goals: predictability, proximity, and diversity. In the diversity selection stage, users can specify the maximum percentage of students eligible for FRL within each zone, as well as bounds on the racial

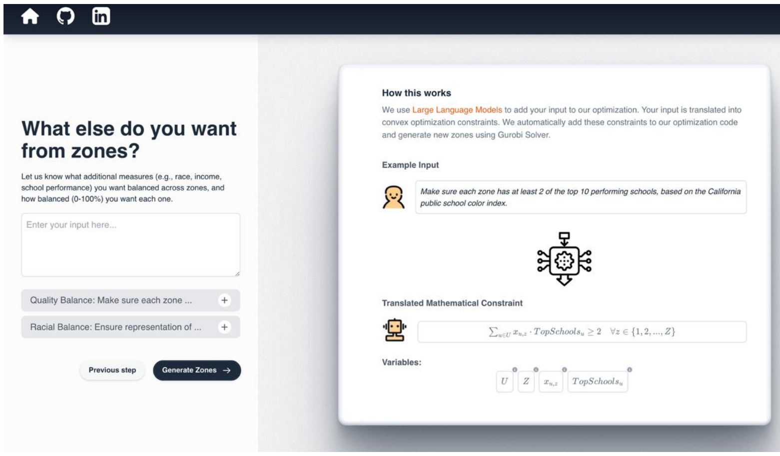


Fig. 5. Interactive constraint exploration tool using language models for zone design

composition of each zone. Users can mark favorite zone configurations and add comments while exploring options. We also provide an extensive dashboard that includes comprehensive metrics requested by stakeholders and map visualizations for each potential zone configuration. We provide screenshots of the dashboard and further details about its use in Appendix B.3.

5 RESULTS

In this section, we describe the results of applying our framework to generate zones and identify promising maps for the San Francisco Unified School District (SFUSD). Our results are illustrative of an ongoing process rather than finalized policy decisions. The iterative elicitation process produced consideration sets of zone maps and quantitative preferences over candidate maps, expressed through scores and elimination constraints. While we focus on evaluating results for one stakeholder group (SFUSD staff), the approach can be used to elicit preferences from other stakeholder groups, and aggregated across multiple stakeholders using standard social choice approaches.

5.1 Data and Current Policy

Each year, around 3,500-5,000 kindergarten applicants participate in the SFUSD student assignment process.[11] They apply to programs at the 72 elementary schools in SFUSD, including nine K-8 programs. In the current student assignment system, students may apply to any program, but receive priority at a single neighborhood school, known as their 'attendance area' school.[12] The proposed policy change will restrict student eligibility for general education programs to within their zone; other programs, including language, special education, and K-8, remain citywide.

Our dataset includes information about kindergarten applicants from 2014-15 to 2022-23, including: home coordinates, demographics (race, special education status, English language learner

[11]Participation in SFUSD student assignment dropped from $\approx$ 5000 kindergarten applicants per year pre-pandemic to $\approx 3500 - 4000$ per year post-pandemic. There was also higher year-to-year variation post-pandemic.

[12] The policy used for student assignment in SFUSD in 2023-24 is a district-wide choice system. Families submit an unrestricted ranked list of preferences for programs across the entire district. Students are then assigned using the Deferred Acceptance algorithm with priorities. The highest priority is given to students with siblings at a program, followed by students living in low-scoring census tracts (CTIP1), and then students from the school's attendance area. A random lottery breaks remaining ties. Further details about this policy can be found in [Allman et al., 2022].

status), and program preferences. We incorporate 2019 free- and reduced-price meal eligibility rates (FRL) by census block, averaged over all grades and students within a block, as an anonymized socio-economic proxy. School-related data includes locations, test scores, and program capacities.

All runs were performed on a 2.6 GHz 6-Core Intel Core i7 with 16GB of RAM. The code was implemented in Python, with the Mixed Integer Programs written using Gurobi. Basic model reductions were applied to reduce the number of variables and constraints when possible (e.g., we do not define decision variables for pairs of units and centroids that are very far apart.)

5.2 Metric Elicitation Results

Through the metric elicitation process, we identified two categories of metrics for zone design. The first category of metrics are baseline constraints that must be satisfied but don't require optimization beyond meeting minimum thresholds. 'School Balance' (3.5) ensures the number of schools in different zones differs by at most one. 'Contiguity' (3.2) ensures all zones are geographically contiguous. 'Overage' ensures the number of unassigned seats in each zone remains below 30%:

$$0.3 \sum_{u \in U} \mathrm{Seat}_u x_{u,z} \geq \sum_{u \in U} (\mathrm{Seat}_u - \mathrm{P}_u) x_{u,z} \quad \forall z \in Z$$

The second category of metrics are those that we actively sought to optimize. 'Shortage (%)' represents the maximum percentage (over all zones) of students within a zone who cannot be assigned a seat, and is given by the smallest value of $\alpha_{capacity}$ such that the following constraints are satisfied:

$$\sum_{u \in U} (\mathrm{P}_u - \mathrm{Seat}_u) x_{u,z} \leq \alpha_{capacity} \sum_{u \in U} \mathrm{P}_u x_{u,z} \quad \forall z \in Z.$$

'FRL Deviation (%)' measures the maximum zone-level deviation from the district's average percentage of students eligible for free or reduced-price lunch, and is given by the smallest value of α_{SES} such that the constraints in (3.7) are satisfied. 'Compactness' is quantified by the number of cut edges (edges between two zones). In addition to optimizing for this measure of compactness, we ultimately relied on stakeholder visual inspection to evaluate zone shapes.

5.3 Constraint Exploration Results

Through the Consideration Set Formation process, stakeholders developed a set of promising positions that included different numbers of zones (6, 10, and 18) across the district. The process revealed key trade-offs: for example, larger zones generally allowed for better balance of both socioeconomic need (measured by FRL Deviation) and shortage of school seats across zones (measured by Shortage). Figure 18 in the appendix illustrates these trade-offs between FRL Deviation (%), Shortage (%), and zone size. The final consideration set consisted of 242 distinct zonings representing 22 different positions. Summary statistics of these positions are provided in Table 2.

Given the consideration set, district staff experimented with the generated solutions using the Constraint Selection dashboard described in Section 4.3. They used a scoring method to choose a set of 25 candidate zone maps from the consideration set of 242 maps. Filtering proceeded in 2 stages. First, a staff member performed a visual review of all maps, eliminating candidates that had impractical compactness failures, or zoning maps divided by highways or natural obstacles (e.g., city parks/hills). Next, maps not in the top quartile of total score, or maps that were not in the top half of the diversity score, proximity score, or predictability score individually were eliminated. Details of the scoring process are provided in Appendix C.2.

5.4 Comparison of Zone-Generation Algorithms

In this section, we evaluate our zone-generation algorithms by comparing the zones they generated for each position selected in Section 5.3. We use our proposed multilevel approach to generate zone maps with 6, 10, 13, 18, and 59 zones. We use our tuned Relaxed ReCom approach to generate zone maps with 6 and 18 zones. Finally, we use our overlapping recursive approach to generate new attendance areas (i.e., 59-zone maps). We compare these to the approach used for policy development in [Allman et al., 2022], which solved the MIP in Section 3.1.1 with attendance areas as units.

Figure 6 compares the performance of zone-generation algorithms for 6-, 18-, and 59-zone maps. Multilevel optimization consistently finds solutions with better zone-level metrics compared to attendance area optimization and Relaxed Recom. Figure 7 shows well-performing 6-zone maps generated by these methods, as well as their zone-level performance on shortage and FRL metrics. Overall, the multilevel approach results in both better zone-level metrics and visually more compact zones. The Relaxed Recom approach does not perform as well, particularly struggling with maintaining compact zones and achieving sufficiently low seat shortage.

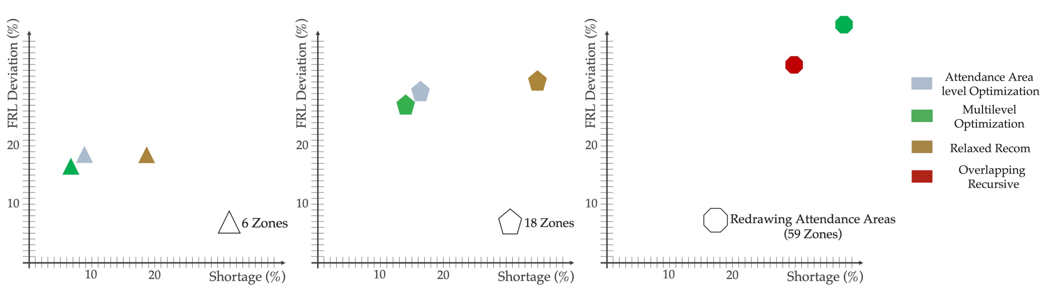


Fig. 6. Comparing zone-generation methods based on their trade-offs between FRL Deviation and Shortage

5.4.1 *Math Programming Approaches.* We find that the effectiveness of the MIP approaches depends on the size of the zones compared to the size of the building blocks. For zone maps with up to 18 zones, the multilevel approach (Algorithm 1) was the most effective in finding solutions satisfying the tightest balance constraints. Some of the resulting zone maps are shown in Figure 8.

Redrawing attendance areas (59 single-school zones) presented unique challenges. The multilevel approach that worked well for larger zones was less effective here, since larger geographic units (current attendance areas, block groups) were a similar size to zones. Instead, we used the overlapping recursive approach (Algorithm 2) at the block level, constructing a zoning map by selecting target and extended school sets and sequencing subproblems. This method required substantial domain knowledge and experimentation to balance efficiency and solution quality, as poorly chosen subproblems could become computationally intractable. The generated attendance areas, shown in Figure 9, significantly reduce variability in socio-economic balance (% FRL) and seat shortage compared to the historical attendance areas.

5.4.2 *Relaxed ReCom.* Figure 10 shows representative examples of applying Relaxed ReCom for 25,000 steps to create 6 zones while balancing multiple constraints. They illustrate common challenges observed across the vast majority of zone maps generated using the Relaxed ReCom approach (see also Figure 19). The resulting zones typically vary significantly in size and shape and are not all compact. Such issues render the solutions impractical for school zoning purposes.

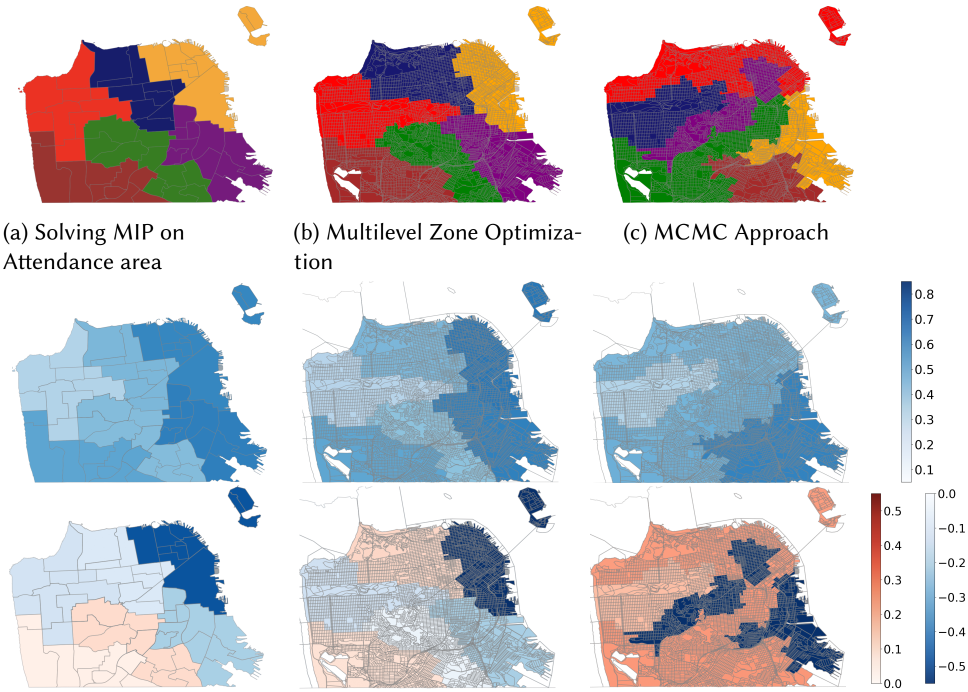
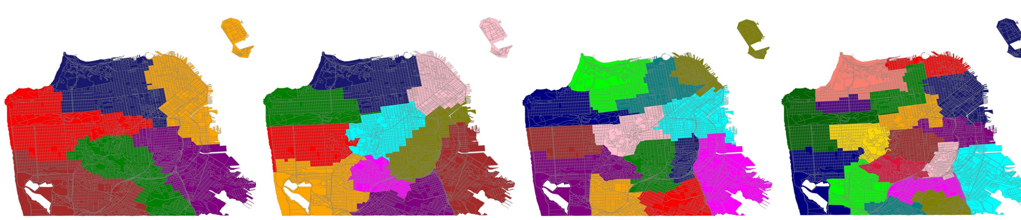


Fig. 7. Comparison of 6-zone solutions generated using different methods. The top row shows zone boundaries, the middle row shows socio-economic balance (%FRL) across zones, and the bottom row shows capacity shortage in each zone (red indicates shortage, blue indicates seat surplus).



(a) 6 Zones (b) 10 Zones (c) 13 Zones (d) 18 Zones

Fig. 8. Zoning solutions with tightest balance constraints for 6, 10, 13, and 18 zones

The challenges encountered by the Relaxed ReCom approach in this setting can be attributed to several factors. The large number of constraints involved in school zoning creates a high-dimensional search space with competing objectives, making it harder for the algorithm to find a solution that balances all the criteria. In Figure 10d, we show one set of zones generated by Relaxed ReCom when only enforcing a single set of balance constraints (i.e., setting $w_{shortage} = w_{Sch} = w_{FRL} = 0$). We see that Relaxed ReCom is able to significantly improve on compactness compared to the multi-constrained problems in Figure 10. However, the presence of irregularly-shaped zones, such as the long tail in the dark blue zone, suggests that even with simpler constraints, the MCMC approach may not be as effective as math programming methods in achieving compactness.

Overall, Relaxed ReCom struggles to navigate the competing objectives effectively, as it aims to satisfy all constraints at once through the biased spanning tree distribution weights and is very

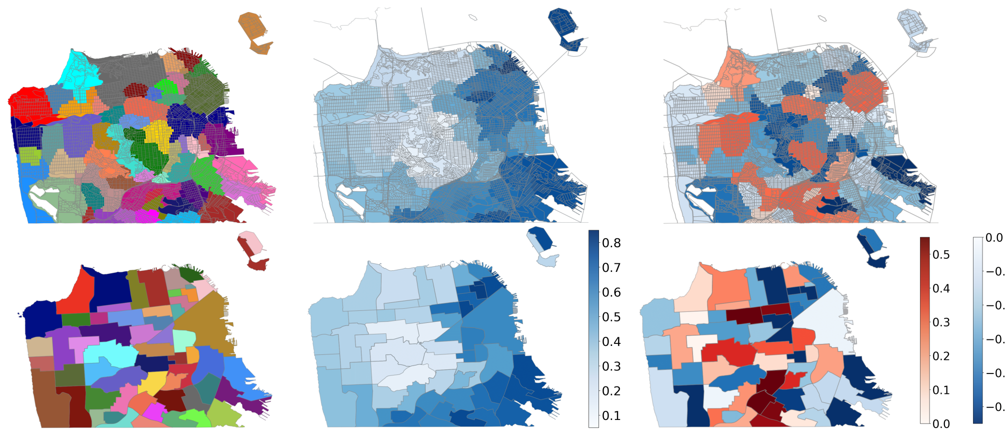
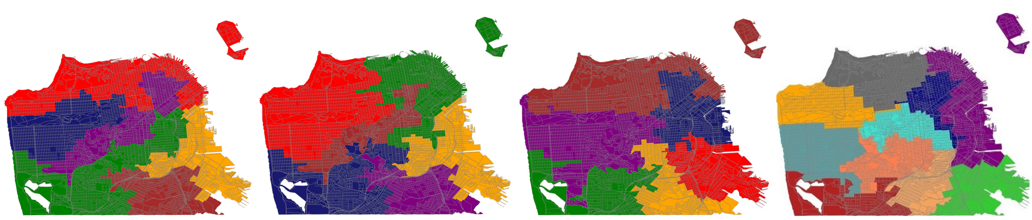


Fig. 9. Comparison of the newly generated 59 attendance areas (top row) and the historical attendance areas (bottom row). The left column shows the zone boundaries for each solution. The middle column illustrates the socio-economic balance (% FRL in each zone). The right column presents the capacity shortage in each zone (red indicates shortage, blue indicates seat surplus).



(a) (b) (c) (d)

Fig. 10. 6-zone maps generated by Relaxed ReCom. The maps in (a)-(c) satisfy all constraints. The map in (d) was found when enforcing only spanning tree balance constraints. Zones are often non-compact, of drastically different sizes (e.g., the yellow and green zones in (a)), and some zones may be stretched or consist of compact areas connected only by a narrow path (e.g., the red zone in (a)).

sensitive to the choice of weights. Attempts to use simulated annealing to explore weights yielded only marginal improvements (Appendix A.8.2). Finding weights that lead to a well-balanced solution across all dimensions is a non-trivial task and may require extensive tuning and experimentation.

5.5 Choice-Driven Evaluation of Selected Zones and Zone-Based Policies

We evaluate several student assignment policies that use our candidate zone maps. Each policy is analyzed using the simulation engine from [Allman et al., 2022], which simulates student preferences over programs within a zone and subsequent assignments using a multinomial logit utility model trained on 2022-23 choice data. We simulate results for kindergarten applicants, programs, and seat capacities in the 2023-24 school year. Results are averaged across 25 simulations.

For the 6, 10, 13, and 18 zone solutions, we evaluate a version of the *Zones+Reserves* policy selected by SFUSD in the 2020 student policy redesign [Allman et al., 2022]. This policy assigns students using the Deferred Acceptance algorithm with zone-based eligibility restriction, and the following priorities and reserves. Students are categorized into two groups based on the percentage of students in their block of residence eligible for free- or reduced-price lunch and reserved a number of seats at each school proportional to their group size. The high- (low-)FRL groups have > 50% (≤ 50%) eligibility in their block. In addition, students with siblings already attending the school are given priority, followed by priority based on a geographical index that takes into account multi-year averages of Level 1 standardized test score rates in the student's area of residence (known as 'CTIP1' priority). Remaining ties are broken by lottery. We also restrict students' general education program eligibility to within their zones. The zone maps for these policies are shown in Figure 8. When evaluating the new attendance areas (59-zone map), we use the district-wide choice policy used by SFUSD in 2023-24, but update the neighborhoods to the attendance areas in Figure 9. We refer to this policy as the New Attendance Area policy, or 'New AA' for short.

We evaluated the resulting assignments based on quantitative metrics capturing the district's priorities of diversity and proximity, as well as choice metrics to assess the disruption to current choice patterns. Table 1 shows the post-assignment metrics for each zoning solution. We compare the selected assignments to a Status Quo policy, which uses the assignment algorithm that was in place in 2023-24 with student preferences given by the estimated choice model.

Table 1. Average assignment metrics

		6 Zones + Reserves	6 Zones + Reserves v2	10 Zones + Reserves	13 Zones + Reserves	18 Zones + Reserves	New AAs	Status Quo
Proximity	*Average Distance (miles)*	1.31	1.32	1.20	1.18	1.19	1.33	1.31
	Distance ≤ 0.5 miles	31.2%	32.2%	35.4%	35.7%	36.5%	34.9%	36.0%
Diversity	*+15% FRL Schools*	11.9	12.2	14.9	14.2	14.7	13.3	12.6
	% AALPI in +15% FRL	15.7%	16.0%	19.9%	21.3%	22.3%	19.8%	18.5%
Choice	*% Top 3 choice*	80.4%	81.2%	80.2%	77.1%	74.9%	87.7%	87.9%

Notes: Preferences are generated by the choice model. Metrics improving on the Status Quo are highlighted in blue.

Metrics: Diversity metrics include the percentage of mid-to-high poverty schools (with ≥ 15% more students eligible for FRL than the district average) and the percentage of historically underserved racial groups (African American, Latinx, or Pacific Islander) assigned to these schools. These metrics are motivated by literature showing the concentration of racial minority students in high-poverty schools is highly correlated with the size and growth of racial academic achievement gaps [Reardon et al., 2024]. Proximity metrics include students' average distance to their assigned school, and percentages of students assigned within a 'walking distance' of 0.5 miles. Choice metrics include the proportion of students assigned one of their top 3 choices. Results for additional metrics can be found in Appendix C.4.

Proximity: All the selected zoning solutions, except the 6-zone policies, improve the average distance to assigned schools compared to the Status Quo. However, only the 18-zone policy increases the percentage of students assigned to a school within 0.5 miles. Smaller zones generally perform better than larger zones in terms of proximity metrics.

Diversity: Larger zones perform better than smaller zones with respect to diversity. Only the 6-Zones+Reserves policies decrease the number of high-poverty schools as well as the percentage of historically underserved ethnic groups in these schools compared to the Status Quo. Smaller zones perform worse on diversity metrics.

Choice: All zone-based policies reduce choice compared to the current policy. This is because they limit students' general education options to schools within their designated zones, and thus may exclude their most preferred programs. However, these metrics improve when we consider assignment to top choices within the zone (Appendix C.4).

Updated Attendance Areas: The newly proposed attendance area boundaries (59 zones) show very similar performance across all metrics compared to the Status Quo.

6 CONCLUSION

Designing multi-school zones for student assignment is a complex task that involves both computational and human challenges. In this paper, we presented a novel approach that draws from interdisciplinary literature on school neighborhood zone design, political redistricting, and participatory design. Our analysis of the San Francisco Unified School District demonstrates that carefully-designed larger zones can effectively enhance both school diversity and student outcomes.

Our work introduced multiple optimization approaches for school zone design, including multi-level and recursive algorithms, which show promise for tackling a wide range of districting problems. While our implementation of MCMC methods for school zoning did not produce usable results for SFUSD, it revealed valuable insights for future research. Understanding the factors contributing to the limitations of the algorithm and exploring alternative biased distributions in the Relaxed ReCom approach could lead to more effective ways of balancing competing objectives.

To address the human dimension of zone design, we developed a framework that pairs optimization techniques with participatory design tools, enabling stakeholders to explore and form their preferences over the feasible set of zone maps and zoning outcomes. This approach extends beyond school districting—it can be applied to any zone-based problem where stakeholders need to understand and evaluate feasible solutions while navigating uncertainty about possible outcomes.

Some of our participatory tools, such as the LLM-driven interactive interface for iterative zone generation, are currently at the prototype stage. We present them here as a proof of concept, demonstrating the potential of such tools to place stakeholders directly in the decision loop rather than relying on a researcher with optimization expertise to translate between natural language and mathematical constraints. Further refinement, user testing with a broad set of stakeholders, and evaluation of these tools in practice remain important directions for future work.

In this paper, we focus on preference elicitation for a single stakeholder. For example, while focus groups involving a wider range of stakeholders were used in the initial metric elicitation stage (Section 4.1), the results from iterative zone-generation reflect only the preferences of SFUSD staff. While the natural next step is to elicit and aggregate preferences from multiple stakeholders, such efforts depend on district timelines for broader community engagement, which remain in flux. Our results nonetheless provide several promising directions for aggregation. Each stakeholder's consideration set can be interpreted as their set of 'acceptable' zones in approval voting [Brams and Fishburn, 2007]. Reasons for eliminating zoning maps can be encoded as constraints, identifying when maps the stakeholder has not seen are unacceptable. Scores from the filtering process can be used to construct stakeholder rank-orderings over consideration sets or zone maps, which can be aggregated via scoring rules such as Borda [Arrow and Raynaud, 1986, Saari, 1995]. Finally, the choice-driven evaluation process can also be used to impute (e.g., via binary comparisons as in [Boutilier et al., 2005]) the relative importance stakeholders place on policy goals, and those weights aggregated into a social welfare function. These aggregation approaches illustrate the flexibility of our framework for multi-stakeholder decision-making. A systematic empirical comparison of aggregation methods is best pursued once the policy implementation timeline allows for collecting richer preference data from a broad range of stakeholders; we leave this for future work.

Overall, our work demonstrates the power of combining rigorous optimization techniques with participatory design to address complex, real-world problems. By bridging the computational and human dimensions, we can develop solutions that, beyond being mathematically optimal, align with the values and priorities of the communities they serve.

REFERENCES

Ali AhmadiTeshnizi, Wenzhi Gao, and Madeleine Udell. 2023. Optimus: Optimization modeling using mip solvers and large language models. *arXiv preprint arXiv:2310.06116* (2023).

Hugo A Akitaya, Andrei Gonczi, Diane L Souvaine, Csaba D Tóth, and Thomas Weighill. 2023. Reconfiguration of Polygonal Subdivisions via Recombination. *arXiv preprint arXiv:2307.00704* (2023).

Maxwell Allman, Itai Ashlagi, Irene Lo, Juliette Love, Katherine Mentzer, Lulabel Ruiz-Setz, and Henry O'Connell. 2022. Designing School Choice for Diversity in the San Francisco Unified School District. In *Proceedings of the 23rd ACM Conference on Economics and Computation*. 290–291.

Nima Anari, Kuikui Liu, Shayan Oveis Gharan, Cynthia Vinzant, and Thuy-Duong Vuong. 2021. Log-concave polynomials IV: approximate exchange, tight mixing times, and near-optimal sampling of forests. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. 408–420.

Ali Aouad and Danny Segev. 2021. Display optimization for vertically differentiated locations under multinomial logit preferences. *Management Science* 67, 6 (2021), 3519–3550.

Kenneth J Arrow and Hervé Raynaud. 1986. Social choice and multicriterion decision-making. *MIT Press Books* 1 (1986).

Eric A Autry, Daniel Carter, Gregory Herschlag, Zach Hunter, and Jonathan C Mattingly. 2008. Multi-scale merge-split markov chain monte carlo for redistricting,(2020). *URL: https://arxiv. org/abs* (2008).

Eric A Autry, Daniel Carter, Gregory J Herschlag, Zach Hunter, and Jonathan C Mattingly. 2021. Metropolized multiscale forest recombination for redistricting. *Multiscale Modeling & Simulation* 19, 4 (2021), 1885–1914.

Sachet Bangia, Christy Vaughn Graves, Gregory Herschlag, Han Sung Kang, Justin Luo, Jonathan C Mattingly, and Robert Ravier. 2017. Redistricting: Drawing the line. *arXiv preprint arXiv:1704.03360* (2017).

Stephen T Barnard and Horst D Simon. 1994. Fast multilevel implementation of recursive spectral bisection for partitioning unstructured problems. *Concurrency: Practice and experience* 6, 2 (1994), 101–117.

Dimitris Bertsimas, Patrick Jaillet, and Sébastien Martin. 2019. Online vehicle routing: The edge of optimization in large-scale applications. *Operations Research* 67, 1 (2019), 143–162.

Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* 2008, 10 (2008), P10008.

Susanne Bødker, Christian Dindler, Ole S Iversen, and Rachel C Smith. 2004. What is participatory design? In *Participatory design*. Springer, 5–13.

Craig Boutilier, Relu Patrascu, Pascal Poupart, and Dale Schuurmans. 2005. Regret-based Utility Elicitation in Constraint-based Decision Problems.. In *IJCAI*, Vol. 9. 929–934.

Steven J. Brams and Peter C. Fishburn. 2007. *Approval Voting*. Springer.

Aydın Buluç, Henning Meyerhenke, Ilya Safro, Peter Sanders, and Christian Schulz. 2016. *Recent advances in graph partitioning*. Springer.

Felipe Caro, Takeshi Shirabe, Monique Guignard, and Andrés Weintraub. 2004. School redistricting: Embedding GIS tools with integer programming. *Journal of the Operational Research Society* 55, 8 (2004), 836–849.

Moses Charikar, Paul Liu, Tianyu Liu, and Thuy-Duong Vuong. 2022. On the complexity of sampling redistricting plans. *arXiv preprint arXiv:2206.04883* (2022).

Ningyuan Chen, Ming Hu, and Wenhao Li. 2022. Algorithmic decision-making safeguarded by human knowledge. *arXiv preprint arXiv:2211.11028* (2022).

Maria Chikina, Alan Frieze, Jonathan C Mattingly, and Wesley Pegden. 2020. Separating effect from significance in markov chain tests. *Statistics and Public Policy* 7, 1 (2020), 101–114.

C. W. Churchman. 1967. Wicked Problems. *Management Science* 14, 4 (1967), B–141–B–146.

Andrew M Davis, Shawn Mankad, Charles J Corbett, and Elena Katok. 2022. The Best of Both Worlds: Machine Learning and Behavioral Science in Operations Management. *Available at SSRN 4258273* (2022).

Daryl DeFord and Moon Duchin. 2019. Redistricting reform in Virginia: Districting criteria in context. *Virginia Policy Review* 12, 2 (2019), 120–146.

Daryl DeFord, Moon Duchin, and Justin Solomon. 2021. Recombination: A family of Markov chains for redistricting. *Harvard Data Science Review* 3, 1 (2021), 3.

Arthur Delarue, Zhen Lian, and Sebastien Martin. 2023. Algorithmic Precision and Human Decision: A Study of Interactive Optimization for School Schedules. *Available at SSRN 4324076* (2023).

Kiera W Dobbs, Douglas M King, and Sheldon H Jacobson. 2023. Redistricting optimization with recombination: A local search case study. *Computers & Operations Research* 160 (2023), 106369.

Thurston Domina, Nikolas Pharris-Ciurej, Andrew M Penner, Emily K Penner, Quentin Brummet, Sonya R Porter, and Tanya Sanabria. 2018. Is free and reduced-price lunch a valid measure of educational disadvantage? *Educational Researcher* 47, 9 (2018), 539–555.

Moon Duchin and Bridget Eileen Tenner. 2024. Discrete geometry for electoral geography. *Political Geography* 109 (2024), 103040.

Benjamin Fifield, Michael Higgins, Kosuke Imai, and Alexander Tarr. 2020. Automated redistricting simulation using Markov chain Monte Carlo. *Journal of Computational and Graphical Statistics* 29, 4 (2020), 715–728.
David Gale and Lloyd S Shapley. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69, 1 (1962), 9–15.
Robert S Garfinkel and George L Nemhauser. 1970. Optimal political districting by implicit enumeration techniques. *Management Science* 16, 8 (1970), B–495.
Sommer Gentry, Eric Chow, Allan Massie, and Dorry Segev. 2015. Gerrymandering for justice: redistricting US liver allocation. *Interfaces* 45, 5 (2015), 462–480.
Nabeel Gillani, Doug Beeferman, Christine Vega-Pourheydarian, Cassandra Overney, Pascal Van Hentenryck, and Deb Roy. 2023. Redrawing attendance boundaries to promote racial and ethnic diversity in elementary schools. *arXiv preprint arXiv:2303.07603* (2023).
Wes Gurnee and David B Shmoys. 2021. Fairmandering: A column generation heuristic for fairness-optimized political districting. In *SIAM Conference on Applied and Computational Discrete Algorithms (ACDA21)*. SIAM, 88–99.
Michael Harwell and Brandon LeBeau. 2010. Student eligibility for a free lunch as an SES measure in education research. *Educational researcher* 39, 2 (2010), 120–131.
John R Hauser and Birger Wernerfelt. 1990. An evaluation cost model of consideration sets. *Journal of consumer research* 16, 4 (1990), 393–408.
Bruce Hendrickson, Robert W Leland, et al. 1995. A Multi-Level Algorithm For Partitioning Graphs. *SC* 95, 28 (1995), 1–14.
Gregory Herschlag, Han Sung Kang, Justin Luo, Christy Vaughn Graves, Sachet Bangia, Robert Ravier, and Jonathan C Mattingly. 2020a. Quantifying gerrymandering in north carolina. *Statistics and Public Policy* 7, 1 (2020), 30–38.
Gregory Herschlag, Jonathan C Mattingly, Matthias Sachs, and Evan Wyse. 2020b. Non-reversible Markov chain Monte Carlo for sampling of districting maps. *arXiv preprint arXiv:2008.07843* (2020).
Gregory Herschlag, Robert Ravier, and Jonathan C Mattingly. 2017. Evaluating partisan gerrymandering in Wisconsin. *arXiv preprint arXiv:1709.01596* (2017).
Charles A Holloway, Donald A Wehrung, Michael P Zeitlin, and Rosser T Nelson. 1975. An interactive procedure for the school boundary problem with declining enrollment. *Operations Research* 23, 2 (1975), 191–206.
Jennifer Jellison Holme. 2002. Buying homes, buying schools: School choice and the social construction of school quality. *Harvard educational review* 72, 2 (2002), 177–206.
John N Hooker and W-J van Hoeve. 2018. Constraint programming and operations research. *Constraints* 23 (2018), 172–195.
Jörg Kalcsics, Stefan Nickel, and Michael Schröder. 2005. Towards a unified territorial design approach—Applications, algorithms and GIS integration. *Top* 13 (2005), 1–56.
George Karypis and Vipin Kumar. 1998. Multilevel algorithms for multi-constraint graph partitioning. In *SC'98: Proceedings of the 1998 ACM/IEEE Conference on Supercomputing*. IEEE, 28–28.
Finn Kensing and Jeanette Blomberg. 1998. Participatory design: Issues and concerns. *Computer supported cooperative work (CSCW)* 7 (1998), 167–185.
Douglas M King, Sheldon H Jacobson, Edward C Sewell, and Wendy K Tam Cho. 2012. Geo-graphs: an efficient model for enforcing contiguity and hole constraints in planar graph partitioning. *Operations Research* 60, 5 (2012), 1213–1228.
Connor Lawless, Jakob Schoeffer, Lindy Le, Kael Rowan, Shilad Sen, Cristina St. Hill, Jina Suh, and Bahareh Sarrafzadeh. 2024. "I Want It That Way": Enabling Interactive Decision Support Using Large Language Models and Constraint Programming. *ACM Transactions on Interactive Intelligent Systems* 14, 3 (2024), 1–33.
Harry A Levin and Sorelle A Friedler. 2019. Automated congressional redistricting. *Journal of Experimental Algorithmics (JEA)* 24 (2019), 1–24.
Karen Levy, Kyla E Chasalow, and Sarah Riley. 2021. Algorithms and decision-making in the public sector. *Annual Review of Law and Social Science* 17 (2021), 309–334.
Cory McCartan and Kosuke Imai. 2023. Sequential Monte Carlo for sampling balanced and compact redistricting plans. *The Annals of Applied Statistics* 17, 4 (2023), 3300–3323.
Anuj Mehrotra, Ellis L Johnson, and George L Nemhauser. 1998. An optimization based heuristic for political districting. *Management Science* 44, 8 (1998), 1100–1114.
Kaisa Miettinen, Jussi Hakanen, and Dmitry Podkopaev. 2016. Interactive nonlinear multiobjective optimization methods. In *Multiple criteria decision analysis: State of the art surveys*. Springer, 927–976.
Tomás E Monarrez. 2023. School Attendance Boundaries and the Segregation of Public Schools in the United States. *American Economic Journal: Applied Economics* 15, 3 (2023), 210–237.
Michael J Muller and Sarah Kuhn. 1993. Participatory design. *Commun. ACM* 36, 6 (1993), 24–28.
Lisa M Nicholson, Sandy J Slater, Jamie F Chriqui, and Frank Chaloupka. 2014. Validating adolescent socioeconomic status: Comparing school free or reduced price lunch with community measures. *Spatial Demography* 2 (2014), 55–65.
Johannes Oehrlein and Jan-Henrik Haunert. 2017. A cutting-plane method for contiguity-constrained spatial aggregation. *Journal of Spatial Information Science* 15 (2017), 89–120.

Cassandra Overney, Cassandra Moe, Alvin Chang, and Nabeel Gillani. 2025. BoundarEase: Fostering Constructive Community Engagement to Inform More Equitable Student Assignment Policies. *Proceedings of the ACM on Human-Computer Interaction* 9, 2 (2025), 1–37.

Ann Owens, Sean F Reardon, and Christopher Jencks. 2016. Income segregation between schools and school districts. *American Educational Research Journal* 53, 4 (2016), 1159–1197.

Aysu Ozel, Karen Smilowitz, and Lila KS Goldstein. 2025. Community-engaged school district design: A stream-based approach. *Operations Research* (2025).

Ariel D Procaccia and Jamie Tucker-Foltz. 2022. Compact redistricting plans have many spanning trees. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM, 3754–3771.

Sean F Reardon, Ericka S Weathers, Erin M Fahle, Heewon Jang, and Demetra Kalogrides. 2024. Is separate still unequal? New evidence on school segregation and racial academic achievement gaps. *American Sociological Review* 89, 6 (2024), 971–1010.

Allison Roda and Amy Stuart Wells. 2013. School choice policies and racial segregation: Where white parents' good intentions, anxiety, and privilege collide. *American Journal of Education* 119, 2 (2013), 261–293.

Francesca Rossi, Peter Van Beek, and Toby Walsh. 2006. *Handbook of constraint programming*. Elsevier.

Alvin E Roth, Uriel G Rothblum, and John H Vande Vate. 1993. Stable matchings, optimal assignments, and linear programming. *Mathematics of operations research* 18, 4 (1993), 803–828.

Donald G/ Saari. 1995. *Basic Geometry of Voting*. Springer.

M Angélica Salazar-Aguilar, Roger Z Ríos-Mercado, and José Luis González-Velarde. 2011. A bi-objective programming model for designing compact and balanced territories in commercial districting. *Transportation Research Part C: Emerging Technologies* 19, 5 (2011), 885–895.

Elizabeth B-N Sanders and Pieter Jan Stappers. 2014. Probes, toolkits and prototypes: three approaches to making in codesigning. *CoDesign* 10, 1 (2014), 5–14.

Satu Elisa Schaeffer. 2007. Graph clustering. *Computer science review* 1, 1 (2007), 27–64.

Peng Shi. 2022. Optimal priority-based allocation mechanisms. *Management Science* 68, 1 (2022), 171–188.

Andrew Spann, Daniel Kane, and Dan Gulotta. 2007. Electoral redistricting with moment of inertia and diminishing halves models. *The UMAP Journal* 28, 3 (2007), 281–299.

Rahul Swamy, Douglas M King, and Sheldon H Jacobson. 2023. Multiobjective optimization for politically fair districting: A scalable multilevel approach. *Operations Research* 71, 2 (2023), 536–562.

Alexis Tsoukiàs. 2008. From decision theory to decision aiding methodology. *European journal of operational research* 187, 1 (2008), 138–161.

Hamidreza Validi and Austin Buchanan. 2022. Political districting to minimize cut edges. *Mathematical Programming Computation* 14, 4 (2022), 623–672.

Marija Vucelja. 2016. Lifting—a nonreversible Markov chain Monte Carlo algorithm. *American Journal of Physics* 84, 12 (2016), 958–968.

Chris Walshaw. 2001. A multilevel algorithm for force-directed graph drawing. In *Graph Drawing: 8th International Symposium, GD 2000 Colonial Williamsburg, VA, USA, September 20–23, 2000 Proceedings 8*. Springer, 171–182.

Chris Walshaw. 2002. A multilevel approach to the travelling salesman problem. *Operations research* 50, 5 (2002), 862–877.

David Bruce Wilson. 1996. Generating random spanning trees more quickly than the cover time. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*. 296–303.

H Peyton Young. 1988. Measuring the compactness of legislative districts. *Legislative Studies Quarterly* (1988), 105–115.

Zhanzhan Zhao, Cyrus Hettle, Swati Gupta, Jonathan Christopher Mattingly, Dana Randall, and Gregory Joseph Herschlag. 2022. Mathematically Quantifying Non-responsiveness of the 2021 Georgia Congressional Districting Plan. In *Equity and Access in Algorithms, Mechanisms, and Optimization*. 1–11.

A COMPUTATIONAL METHODS

A.1 Combining Zone Optimization with Choice

In order to combine zone optimization and choice into a single, tractable math program, we use a priority optimization formulation similar to [Shi, 2022].

Let S be the set of schools, and let T be the set of student types, where students of type $t \in T$ have utilities drawn from distribution F_t. Let $n_{u,t}$ be the number of students in unit u of type t $\forall u \in U, t \in T$, and let $\text{Util}_t(S') := \mathbb{E}_{f \sim F_t}[\max_{u \in S} f_u]$ be the expected utility of a type t student with budget set $S' \subseteq S$. We further define $P_t(s, S') := 1\{s \in S'\}\mathbb{P}_{f \sim F_t}(f_s = \max_{s' \in S'} f_{s'})$ is the probability a student of type t's favorite school in S' is s. Let variable $y_{t,u,S'}$ be the probability that a student of type t living in unit u has budget set S'.

$$
\begin{aligned}
\max_{y} \quad & \sum_{u,t,S'} n_{u,t}\text{Util}_t(S')y_{t,u,S'} & & \\
\text{s.t.} \quad & & y_{t,u,S'} \geq 0 \quad & \forall u \in U, t \in T, S' \subseteq S \\
& \text{(Capacity)} & \sum_{t,u,S'} n_{u,t}P_t(s, S')y_{t,u,S'} \leq c_s \quad & \forall u \in U, t \in T, s \in S \\
& \text{(Valid probabilities)} & \sum_{S' \subseteq S} y_{t,u,S'} = 1 \quad & \forall u \in U, t \in T
\end{aligned}
\tag{P}
$$

A.2 Combining zone optimization and priority optimization and choice.

Suppose we want to combine zones with priority optimization and choice by directly using math program (P). Then we need to introduce additional constraints on valid probabilities $y_{t,u,S'}$. Specifically, we need to ensure that a student's budget set is contained in their zone, i.e.,

$$y_{t,u,S'} > 0 \text{ only if } S' \subseteq \{\text{schools in } u\text{'s zone}\}.$$

A school s is in student u's zone if and only if $\sum_{z \in Z} x_{s,z}x_{u,z} = 1$, and otherwise $\sum_{z \in Z} x_{s,z}x_{u,z} = 0$. So we can ensure each students' budget set is contained in their zone using the *quadratic* constraints

$$y_{t,u,S'} \leq \sum_{z \in Z} x_{s,z}x_{u,z} \; \forall t \in T, S' \subseteq S, s \in S. \tag{A.1}$$

Making the quadratic constraints linear. In this subsection, we provide an alternative formulation of the zone optimization program that increases the number of variables and constraints (by a factor of $|Z|/|S|$) but allows us to characterize valid probabilities using a linear constraint. Let $x_{u,s}$ be a decision variable indicating whether unit u is part of the same zone as school s, defined for all $u \in U$, $s \in S$. We write the following optimization program.

$$
\begin{aligned}
\min \quad & \sum_{u,t,S'} n_{u,t} F_t(S') y_{t,u,S'} + \sum_{u\in U,\, v\in N(u)} b_{u,v} && && \text{(A.2)}\\
\text{s.t.} \quad & |x_{u,s} - x_{v,s}| \leq 2 - (x_{u,s'} + x_{v,s'}) && \forall v,u \in U, s,s' \in S && \text{(A.3)}\\
& y_{t,u,S'} \leq x_{u,s} && \forall u \in U, t \in T, S' \subseteq S, s \in S' && \text{(A.4)}\\
& y_{t,u,S'} \geq 0 && \forall u \in U, t \in T, S' \subseteq S && \text{(A.5)}\\
& \sum_{t,u,S'} n_{u,t} P_t(s,S') y_{t,u,S'} \leq c_s && \forall u \in U, t \in T, s \in S && \text{(A.6)}\\
& \sum_{S'\subseteq S} y_{t,u,S'} = 1 && \forall u \in U, t \in T. && \text{(A.7)}\\
& x_{u,s} \leq \sum_{v: v\in N(u), d_{v,s} \leq d_{u,s}} x_{v,s} && \forall u \in U, s \in S && \text{(A.8)}\\
& 0.15 \cdot \sum_{u\in U} n_u x_{u,s} \geq |\sum_{u\in U} (n_u - q_u) x_{u,s}| && \forall s \in S && \text{(A.9)}\\
& 1 \geq |\sum_{u\in U} \frac{sch_u}{|Z|} - \sum_{u\in U} sch_u x_{u,s}| && \forall s \in S && \text{(A.10)}\\
& \sum_{u\in U} f_u x_{u,s} \geq (\frac{F}{N} - 0.15) \cdot \sum_{u\in U} n_u x_{u,s} && \forall s \in S && \text{(A.11)}\\
& \sum_{u\in U} R_u^k x_{u,s} \geq (\frac{R^k}{N} - 0.15) \cdot \sum_{u\in U} n_u x_{u,z} && \forall s \in S, k \in K && \text{(A.12)}
\end{aligned}
$$

Note that this program is almost identical to combining the previous two math programs and replacing all instances of $x_{u,z}$ with $x_{u,s}$. There are a few notable differences, which are as follows. In the objective, we have replaced the compactness objective with a choice objective (total utility). Constraint (A.3) is the zone constraint. It ensures that if two units u, v are in the same zone as some school s', then the set of schools in u's zone is the same as the set of schools in v's zone (i.e. $x_{u,s} = x_{v,s}$). Otherwise, the units do not share any schools. Constraint (A.4) is the constraint for valid probabilities. It ensures that a student's budget set is contained in their zone. Constraints (A.5)-(A.7) are as in the math program (P) in subsection A.1, and all other constraints (A.8)-(A.12) are as in the math program in subsection 3.1, with $x_{u,z}$ replaced with $x_{u,s}$ for all $s \in S$ rather than for all $z \in Z$.

A.3 Endogenous Centroids

In this subsection, we provide an integer program in which centroids are variables. This allows our integer program to find a larger class of solutions, and potentially including zonings satisfying constraint sets which we previously deemed infeasible. Unfortunately, this integer program has substantially more variables, which makes it extremely computationally expensive to solve.

Let y_z be a binary decision variable indicating whether unit z is the centroid of a zone or not, $\forall z \in S$. Let $x_{u,z}$ be the decision variable indicating whether unit u is part of a zone with unit z as its centroid-unit or not, $\forall u, z \in U$. In this new integer program, we enforce all the previously introduced constraints, but anywhere among constraints (3.1)-(3.8) where we had the term "$z \in Z$" we now replace it with "$z \in U$". We additionally enforce the following constraints.

$$\sum_{z \in U} y_z = |Z| \tag{A.13}$$

Constraint (A.13) makes sure we have exactly $|Z|$ centroids.

$$x_{u,z} \leq y_z \qquad \forall u, z \in U \tag{A.14}$$

Constraint (A.14) makes sure we only assign a unit u to a zone with centroid z if unit z is considered as a centroid.

A.4 Contiguity Constraint

Figure 11 illustrates a case where the contiguity constraint (3.2) may fail: unit a does not have any neighbors closer to centroid z than itself. To address this issue, we enforce constraint (3.2) only for units u and zones z where u has at least one neighbor closer to z.

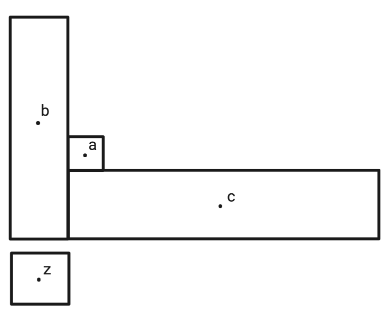


Fig. 11. Visualizing the Limitations of the Contiguity Constraint in (3.2)

A.5 Local Search

Starting with the solution obtained from the Multilevel Zone Design process (Algorithm 1) in Section 3.1.2 , we apply the following Local Search algorithm (Algorithm 4) to further improve the zoning results. In each step, the Local Search algorithm randomly selects two adjacent zones, and merges and splits them into two new zones that minimize boundary costs while satisfying MIP constraints. The local search process is repeated for a maximum of 5000 iterations or until convergence is reached (whichever occurs first). It is important to note that the Local Search algorithm is designed to operate at the census block level for the geographical units. This is possible because at each iteration, the algorithm only needs to solve a bipartition problem on a subset of the district.

ALGORITHM 4: Local Search

Input: Current school zoning, P. Let $P(u)$ be the zone for unit u in zoning P.
Output: New school zoning, Q.
– Randomly select two adjacent zones Z_1 and Z_2 in P.
– Merge zones Z_1 and Z_2 to create a single partition U.
– Split the partition U into two zones, $\hat{Z}_1$ and $\hat{Z}_2$, that minimize boundary costs while satisfying MIP constraints.
– Define $Q(u) = \begin{cases} \hat{Z}_1 & \text{if } u \in \hat{Z}_1 \\ \hat{Z}_2 & \text{if } u \in \hat{Z}_2 \\ P(u) & \text{otherwise} \end{cases}$
Return Q

A.6 Divide-and-Conquer Zoning

Motivated by the non-overlapping recursive zoning algorithms in [Gurnee and Shmoys, 2021, Levin and Friedler, 2019], we explored the following recursive algorithm that divides the district into non-overlapping parts and solves each subproblem separately (Algorithm5, Figure 12).

ALGORITHM 5: Divide and Conquer

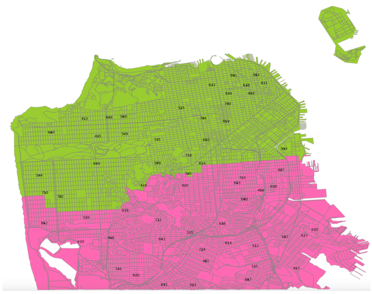
Divide: Divide the district into two balanced parts, A and B, minimizing the shortage of school seats and boundary length (Figure ref).
Recursively divide: At step i, recursively sub-divide each part into 2^i balanced sections, minimizing boundary costs while ensuring an equal number of schools per section and balancing shortage of school seats across sections.
Improve boundaries: Improve boundaries by: trimming (i.e., discarding) zone assignment for any units with few adjacent units in the same zone;[13] and then reassign trimmed units to zones to minimize boundary costs while ensuring an equal number of schools per section and balancing shortage of school seats across sections.

Notes: In the recursive dividing step, we experimented with different splitting patterns (e.g., 2/4/8, 4/2/8) but focus here on the 2/4/8 split for illustrative purposes. When improving boundaries, for SFUSD, we first trim the units which have over 90% of their neighboring units in different zones. We then trim units which (after the initial trimming) lack a contiguous path to their zone's centroid.

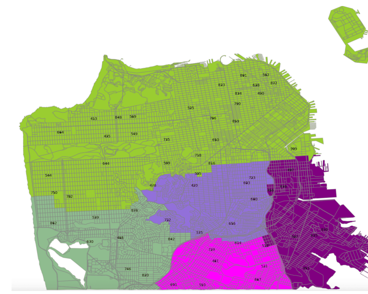
One limitation of the Divide and Conquer approach is that the zone corners or edges become fixed due to previous partitioning steps when the parts (A_i or B_i) are divided into smaller zones. This restricts flexibility in adjusting zone boundaries and limits solution space variation (Figure 12d). As a result, this approach struggles to achieve a good balance across zones in terms of shortage and socio-economic diversity. In addition, a solution that appears compact and well-balanced when dividing the district into two parts may not be optimal when further subdividing parts into smaller sections. Initial partitioning decisions significantly impact the quality of the final zoning solution. Solutions obtained using Divide and Conquer have significantly worse balance compared to the Overlapping Recursive Zoning introduced in Section 3.1.3. For example, zones may have a shortage of up to 45%, much higher than the 28% shortage achieved by solutions in Section 3.1.3.

A.7 Jointly Optimizing Resource Allocation and Zone Design

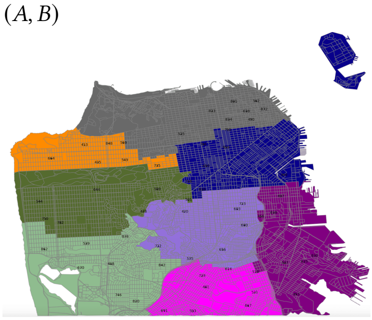
In this section, we propose a solution to optimize the allocation of educational programs across schools in San Francisco while maintaining the total number of program offerings. The current fixed



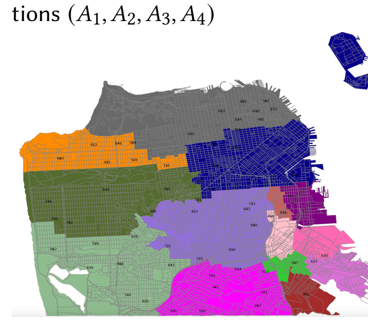
(a) Dividing the map into two parts, (A, B)



(b) Sub-divide part A into four new sections (A_1, A_2, A_3, A_4)



(c) Sub-divide part B into four new sections (B_1, B_2, B_3, B_4)



(d) Sub-divide part (B_1) into 8 new sections

Fig. 12. Iterations of the Divide and Conquer process

locations of schools pose a challenge, as the existing distribution of schools and their capacities may not optimally serve the student population. Altering these locations would be prohibitively expensive due to the high costs associated with closing and opening schools. However, while the physical school locations cannot be easily changed, we suggest that modifying the allocation of classrooms within each school can better meet the needs of students.

In SFUSD, kindergarten students apply to 30 special education programs (8 different types) that are crucial for supporting students with disabilities. In the 2019-20 academic year, over 3,400 children in grades K-5 were enrolled in these programs, with 78% identified for Speech or Language Impairment, Specific Learning Disability, Autism, and Other Health Impairment. Due to the wide variety and limited availability of special education programs, the proposed zoning system allows students to be eligible for all such programs district-wide, while restricting their eligibility for GE programs to their assigned zone.

Despite the importance of these programs, several equity gaps exist. Students enrolling in special education programs are more likely to come from marginalized communities. Suboptimal program placement has led to increased travel times and transportation costs for students, and enrollment patterns reveal disparities in access to different Special Education programs based on disability type and race.

SFUSD is obligated to provide transportation for students enrolling in special education programs due to their specific needs and the limited number of programs available across the district. The cost of busing for special education is substantial, amounting to $26 million per year, primarily because students do not always live close to the schools offering the programs they require. However, the demand for special education programs is not entirely unpredictable, as specific neighborhoods have historically shown consistent demand for particular programs over time.

Proposed Solution. To address these challenges, we propose reallocating programs across schools while maintaining the total number of offerings for each program. By offering programs at schools located close to areas with high demand, we aim to reduce transportation costs. The new problem simultaneously tackles two goals:

- Find a relocation of special education programs that minimizes the total cost of transportation for students enrolling in these programs.
- Divide the district into a specified number of zones, limiting general education (GE) students' choices to schools within their assigned zone.

Incorporating post-choice outcomes, such as transportation costs for special education students, directly into the objective function is not feasible (as shown in Appendix A.1). Instead, we model the problem as a centralized system that assigns students to special education programs while minimizing total transportation cost. This heuristic provides a good estimate of transportation costs, guiding the effective relocation of programs.

We aim to develop a single optimization model that addresses both Goal 1 (relocating classes) and Goal 2 (finding the best zoning). A combined approach can lead to better solutions that might be missed in a two-step optimization heuristic. By simultaneously relocating programs and finding the best assignment of students to schools for special education, we can achieve a better results.

Let T be the set of special education programs, including Mild/Moderate Autism Focus (AF), Emotionally Disturbed (ED), Spanish Immersion (SE), etc. Let $\text{Distance}_{u,s}$ be the transit distance between unit u and school s. Let Population_u^t be the number of students in unit u that would like enroll in program t. Let $A_{u,s}^t$ be a binary variable indicating whether students of Population_u^t are assigned to school s. To generate zones for GE students while optimizing the transit distance for students assigned to special education programs, we modify the mixed integer program from subsection 3.1.

Objective. The objective function is given by

$$Obj_{SpEd} = \sum_{t\in T}\sum_{u\in U}\sum_{s\in S} \text{Distance}_{u,s} \cdot \text{Population}_u^t \cdot A_{u,s}^t.$$

It computes the total transit distance of each special education program student to their matched school, which will be combined with the previous objective value 3.3 for generating zones for GE students. This way we simultaneously optimize the transit distance of students assigned to relocated special education programs classes while finding compact zones for GE students.

Assignment Feasibility. The feasibility constraint (A.15) ensures that students of all special education programs, from all neighborhoods, are assigned to a school that offers that program.

$$\sum_{s\in S} A_{u,s}^t = 1 \qquad \forall t \in T, u \in U \tag{A.15}$$

Assignment Capacity. In SFUSD, each school can have a maximum of only one class for each special education program. This ensures that the limited number of classes for each program is distributed across different schools rather than concentrated in a single school. In contrast, schools are allowed to have multiple classrooms for general education (GE). To model this specific criterion, we introduce the decision variable q_s^t, which indicates whether school s has a class for students of program t. Let Capacity^t be the capacity of a classroom for program $t \in T$, considering the required equipment and the number of students a teacher can handle for each specialty need.

$$\sum_{u\in U} \text{Population}_u^t \cdot A_{u,s}^t \leq \text{Capacity}^t.q_s^t \qquad \forall t \in T, s \in S \tag{A.16}$$

Constraint (A.16) ensures that for each school s, there are enough seats for all the students with each special education type matched to school s.

Modeling the classrooms for GE is slightly more complex since schools can have multiple GE classrooms. We introduce a separate variable for each GE class in each school. Let $y_{s,z}^i$ be a binary variable indicating whether the ith classroom of school s is offering GE program, and is part of zone z. Let Rooms_s be the number of classrooms in school s.

$$y_{s,z}^i \geq y_{s,z}^{i+1} \qquad \forall s \in S, z \in Z, i < \text{Rooms}_s \tag{A.17}$$

Constraint (A.17) ensures that the GE classes in each school are assigned to only one zone, preventing the optimization model from assigning different GE classrooms of a single school to separate zones.

Zone Alignment. In this novel formulation, both classes and geographic units (containing schools) are assigned to zones. We introduce constraint (A.18) to ensure that the classes of school s are assigned to zone z only if the geographic unit in which school s is located is also assigned to the same zone:

$$x_{s,z} \geq y_{s,z}^1 \qquad \forall s \in S, z \in Z \tag{A.18}$$

In the zone optimization model in subsection 3.1, we assigned each unit to a zone using decision variables $x_{u,z}$. Now, our GE classrooms within a school must follow the same logic: a school with its GE classes can be assigned to a zone only if that unit area is assigned to that zone. It is sufficient to form the inequality constraint for $y_{s,z}^1$, given constraint (A.17).

Note: The introduction of the variables $y_{s,z}^i$ might seem unnecessarily complicated, but they are crucial for finding the zoning solution simultaneously using a linear model. If we were to use a simpler variable q_s^{GE} to represent the number of GE classes in a school s, similar to the approach used for special education classes, we would encounter a quadratic constraint when computing the shortage of each zone. This constraint would limit the value of $\sum_{u\in U} x_{u,z}(q_u^{GE} - \text{Population}_u^{GE})$, which is a product of two decision variables. Quadratic constraints are much more computationally expensive to solve, even though the model would have fewer variables compared to our linear formulation using $y_{s,z}^i$.

Unchanging Classroom Quotas. Although we allow programs to be moved to different schools, we want to ensure that the total number of programs in each school, and the total number of programs of each type, stay the same.

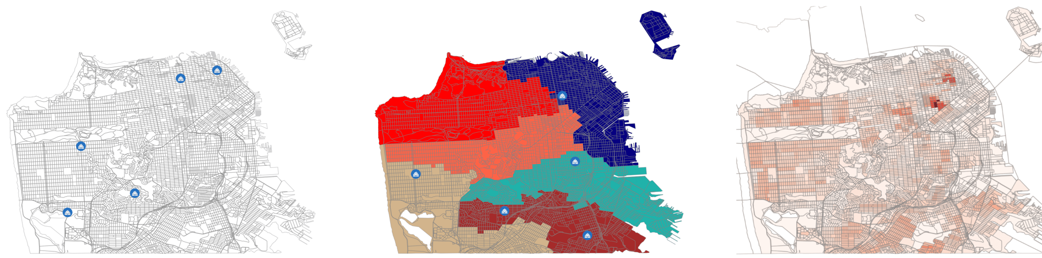
$$\text{Rooms}_s = \sum_{t\in T} q_s^t + \sum_{z\in Z}\sum_i y_{s,z}^i \qquad \forall s \in S \tag{A.19}$$

Constraint (A.19) ensures that the total number of classes in each school remains the same, as the physical limits on the number of rooms in the building prevent increasing or decreasing the total number of classes within a school.

Let Classes^t be the total number of classes offered for program t across the district.

$$\text{Classes}^t = \sum_{s\in S} q_s^t \qquad \forall t \in T \tag{A.20}$$

Constraint (A.20) ensures that the total number of classes for each special education program remains constant across the district. At this phase of the work, we do not want to increase or decrease the overall number of programs offered in the district.



(a) Location of Autism Focus (AF) Programs in SF.

(b) New Location of AF Programs in SF, along with new set of zones.

(c) Distribution of students eligible for AF programs across SF from 2014-2023

Fig. 13. Using MIP to relocate Autism Focus (AF) Programs.

Shortage. To generate zones for GE students while optimizing the transit distance for students enrolling in special education programs, we modify the mixed integer program from subsection 3.1. All constraints remain the same as in subsection 3.1, with the exception of the shortage constraint (3.6), which is redefined as follows.

$$0.15 \cdot \sum_{u \in U} x_{u,z}.\text{Population}_u^{GE} \geq \left| \sum_{u \in U} \text{Population}_u^{GE} - \text{Capacity}^{GE}. \sum_{z \in Z} \sum_{i} y^i_{s,z} \right| \qquad \forall z \in Z \qquad \text{(A.21)}$$

Constraint (A.21) ensures that for each zone z, there are enough seats for GE students assigned to zone z among all the GE classes assigned to that zone:

Finally, we have the following constraints on the decision variables:

$$y^i_{s,z}, q^t_s, A^t_{u,s} \in \{0, 1\} \qquad \forall u \in U, z \in Z, s \in S, i < \text{Rooms}_s$$

By incorporating these constraints and objectives into a single optimization model, we can effectively reallocate special education programs across schools in San Francisco while simultaneously drawing optimal zones for general education students. This approach aims to reduce transportation costs for special education programs and improve the overall efficiency of the school system.

Results. We analyze the change in the distribution of the Autism Focus (AF) program, based on our optimization model.

Figure 13a illustrates the current location of AF programs in the San Francisco Unified School District. To optimize the distribution of these programs, we run our model with the objective of reducing transit times for AF students while simultaneously reallocating school classes and drawing zone boundaries to divide the district into 6 zones (used here as a sample example). The results of this optimization are presented in Figure 13b.

To further demonstrate the effectiveness of our approach, we analyze the distribution of students who participated in AF programs from 2014-2023. The optimized locations of AF programs exhibit a higher proximity to areas with a high concentration of students eligible for AF programs. Moreover, the new AF program locations are not only closer to these students but also more evenly distributed throughout the district. These improvements in proximity and distribution contribute to reduced transit times for students enrolled in AF programs, ultimately enhancing the accessibility and efficiency of the special education program.

A.8 Other Computational Approaches

A.8.1 Constraint Programming. Constraint Programming (CP) is a powerful paradigm for solving combinatorial problems, using both arithmetic and logical algebra to handle highly disjunctive (non-linear) constraints [Rossi et al., 2006]. CP has been applied to a wide range of domains, including scheduling, planning [Hooker and van Hoeve, 2018], and vehicle routing [Bertsimas et al., 2019]. [Gillani et al., 2023] apply constraint programming to redraw attendance boundaries around each elementary school, aiming to reduce school segregation.

We implemented a Constraint Programming approach formulated in the CP SAT-solver. This approach can speed up computation by a factor of 3 compared to the direct mixed-integer programming approach. This speedup was insufficient for CP to directly solve the zoning problem at the scale required for SFUSD. However, it suggests that a multilevel or recursive constraint programming approach may hold promise. This direction is especially promising given recent advances in providing interactive decision support by combining LLMs with constraint programming [Lawless et al., 2024].

A.8.2 Simulated Annealing. Simulated annealing allows the algorithm to escape local optima by occasionally accepting worse solutions, especially early on when the 'temperature' is high [see, e.g., Fifield et al., 2020, Herschlag et al., 2017, for applications to MCMC redistricting]. To address the challenges faced by Relaxed ReCom, we applied simulated annealing to our problem. Starting with power weights of 0 in the biased spanning tree distribution, we sampled for 5,000 steps, then gradually increased the weights to their final values over 10,000 steps, and maintained the final weights for an additional 5,000 steps before taking a sample. The solutions obtained showed only marginal improvements in balancing metrics across zones.

B ADDITIONAL DETAILS FOR SECTION 4

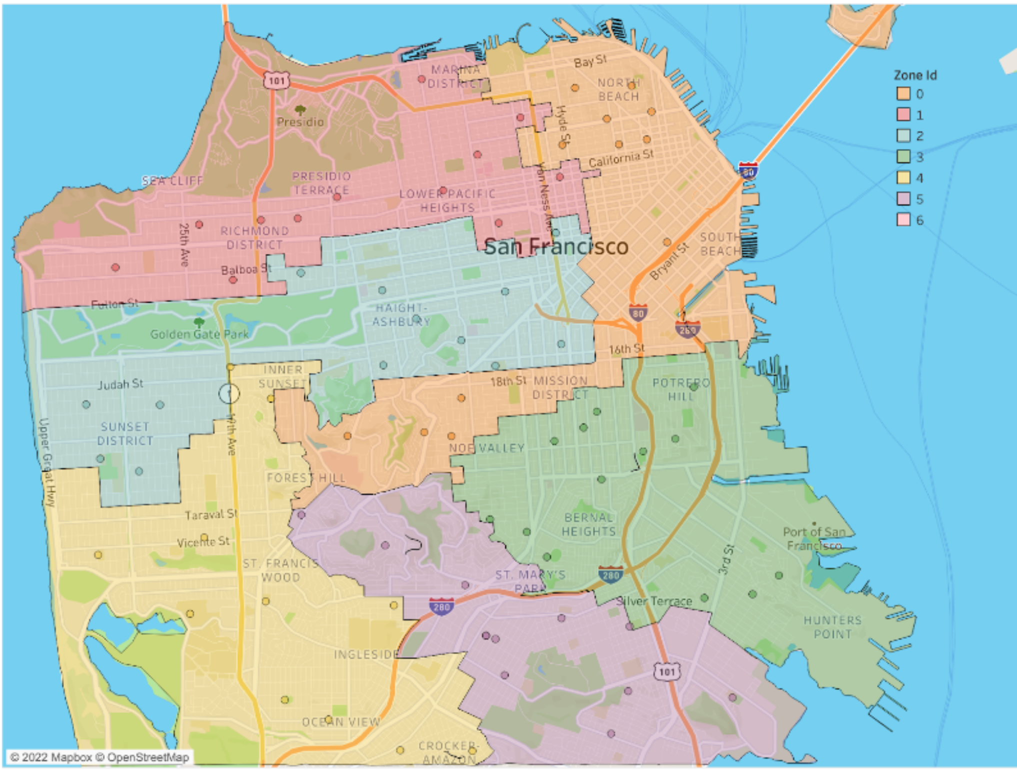
B.1 Focus Groups

The questions used in the focus groups were as follows.

(1) Look at the shape of the zones.
  - Does anything look good about them?
  - Does anything raise concerns?
  - Are there any zones that really fit the geography of the city (hills, highways, parks, etc.)?
  - Any zones that do not make geographic sense?

(2) Using your expertise as a member of your community, neighborhood, or organization, how would the zones affect you and your community?
(3) Are there any absolute deal breakers on this map? What do you feel is non-negotiable?
(4) What creative solutions can you share for any problems you see?
(5) What other feedback do you have about this map? What do you want to see from future maps to better serve students and families?

Leaders provided participants with a sample zone map, and walked through sample discussion prompts to ensure participants understood the task. Then, participants were instructed to look at groupings of 3 zone maps. In addition to the 3 maps, there were brief metrics capturing socio-economic diversity, travel times, and school performance. Figure 14 shows an example of this



Map scores out of 10

The higher the score, the better this map promotes our policy goals

Total map score: 7.1

Diversity score: 7.5

Predictability score: 7.3

Proximity score: 5.6

Map Total Score

5.8 7.2

Table Name

GE_Zoning_2_1_1_2_2_centroids 6-zone-9_BG_ls.shp

☐ Show history

Primary Metrics

Zone Id	% of students who qualify for free or reduced meals	Average home to school walking time (minutes)	Average school performance score (out of 5.00)
0	58.2%	62.33	3.68
1	40.9%	60.97	3.80
2	40.6%	69.42	3.73
3	58.0%	46.44	2.79
4	46.0%	62.80	3.48
5	55.9%	56.16	3.13
6			
Grand Total	50.1%	59.13	3.41

Zone Id: 0, 1, 2, 3, 4, 5, 6

Fig. 14. Example map and metrics used in zone design focus groups.

map and accompanying metrics. Participants then had unstructured discussion time to investigate the maps. They were encouraged to think out loud, and leaders largely tried to remain out of the discussion, only asking clarifying questions. After roughly 10 minutes of discussion, the group would move on to a second set of maps. The discussion process was repeated 2-3 times, showing participants a total of 6 to 9 maps. The specific maps shown were selected by district staff to represent the broadest possible zone design possibilities. After moving through each group of maps, participants and leaders then regrouped for a final reflection on the zone exploration process.

B.2 Illustrations of Sequential Consideration Set Formation (Section 4.2)

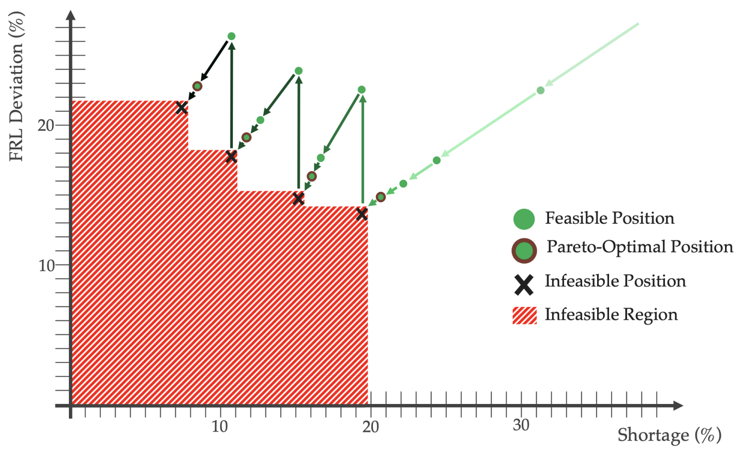


Fig. 15. Stakeholders' search process.

Figure 15 illustrates our iterative constraint discovery process and stakeholder interactions. In this example, the stakeholder starts at feasible positions with relatively loose constraints (e.g., obtained through an earlier zone generation process). They iteratively seek to tighten both socio-economic balance and shortage constraints, and when they reach an infeasible position loosen the socio-economic balance constraint. Green circles indicate feasible positions that are identified by the search process in this example, and outlined green circles indicate Pareto-optimal positions (amongst those in the consideration set). Crosses indicate infeasible positions, from which it can be inferred that the entire shaded region is infeasible.

While the figure shows a single request trajectory in a 2D map of socio-economic balance and shortage, our actual process involves multiple parallel requests, each exploring various directions. In practice, we consider additional metrics, including racial distribution (e.g., the proportion of racial minorities in each zone) and the inclusion or exclusion of K-8 schools in zones.

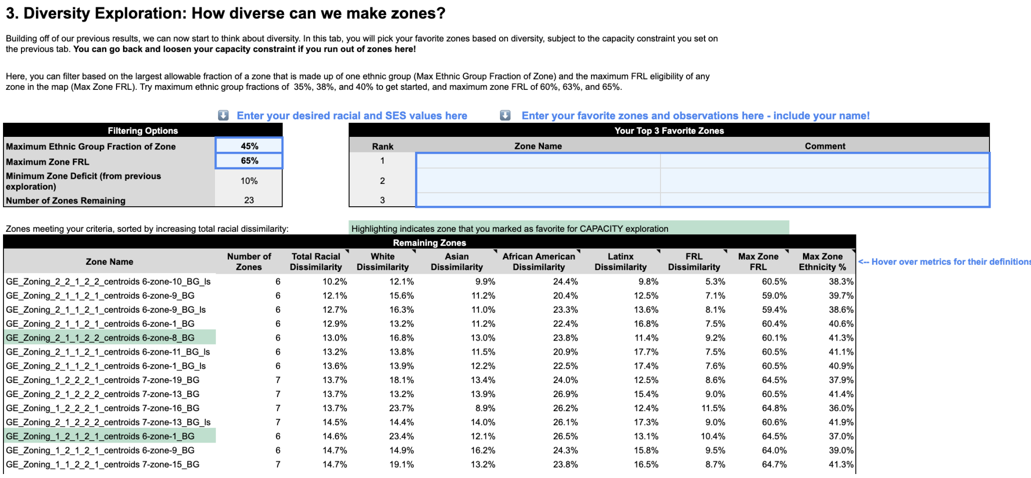
B.3 Interactive Selection Tool

The interactive selection tool was developed using Google Sheets, a familiar and easy-to-use platform that allows for building interactive web interfaces without extensive web development experience. The choice of platform was inspired by the success of a similar tool used by SFUSD decision-makers to explore trade-offs in school start times [Delarue et al., 2023].

The tool incorporates several design features to enhance user experience, such as color coding, tooltips, expandable sections for detailed metrics, and a plot comparing the current zone's performance to other zones in the solution family (Figure 17). Inspired by gerrymandering identification strategies, the tool highlights the selected map's position in a histogram of metrics like racial dissimilarity, illustrating the trade-offs made relative to other maps in the collection. These design choices aim to simplify the complex task of comparing maps.

The tool was used by a small number of district staff with a high degree of data literacy and familiarity with the student assignment problem to select a small number of well-performing zones for further discussion. Their process is described in Section 5.3. The selected zones include those shown in Figure 8, which we evaluated in Section 5.5.

Initial testing with a wider set of district staff revealed that the tool was too complicated for the vast majority of participants due to participants' insufficient familiarity with the problem and



3. Diversity Exploration: How diverse can we make zones?

Building off of our previous results, we can now start to think about diversity. In this tab, you will pick your favorite zones based on diversity, subject to the capacity constraint you set on the previous tab. **You can go back and loosen your capacity constraint if you run out of zones here!**

Here, you can filter based on the largest allowable fraction of a zone that is made up of one ethnic group (Max Ethnic Group Fraction of Zone) and the maximum FRL eligibility of any zone in the map (Max Zone FRL). Try maximum ethnic group fractions of 35%, 38%, and 40% to get started, and maximum zone FRL of 60%, 63%, and 65%.

Enter your desired racial and SES values here

Filtering Options	
Maximum Ethnic Group Fraction of Zone	45%
Maximum Zone FRL	65%
Minimum Zone Deficit (from previous exploration)	10%
Number of Zones Remaining	23

Enter your favorite zones and observations here - include your name!

Your Top 3 Favorite Zones		
Rank	Zone Name	Comment
1		
2		
3		

Zones meeting your criteria, sorted by increasing total racial dissimilarity:

Highlighting indicates zone that you marked as favorite for CAPACITY exploration

<-- Hover over metrics for their definitions!

Remaining Zones									
Zone Name	Number of Zones	Total Racial Dissimilarity	White Dissimilarity	Asian Dissimilarity	African American Dissimilarity	Latinx Dissimilarity	FRL Dissimilarity	Max Zone FRL	Max Zone Ethnicity %
GE_Zoning_2_2_1_2_2_centroids 6-zone-10_BG_ls	6	10.2%	12.1%	9.9%	24.4%	9.8%	5.3%	60.5%	38.3%
GE_Zoning_2_1_1_2_1_centroids 6-zone-9_BG	6	12.1%	15.6%	11.2%	20.4%	12.5%	7.1%	59.0%	39.7%
GE_Zoning_2_1_1_2_1_centroids 6-zone-9_BG_ls	6	12.7%	16.3%	11.0%	23.3%	13.6%	8.1%	59.4%	38.6%
GE_Zoning_2_1_1_2_1_centroids 6-zone-1_BG	6	12.9%	13.2%	11.2%	22.4%	16.8%	7.5%	60.4%	40.6%
GE_Zoning_2_1_1_2_2_centroids 6-zone-8_BG	6	13.0%	16.8%	13.0%	23.8%	11.4%	9.2%	60.1%	41.3%
GE_Zoning_2_1_1_2_1_centroids 6-zone-11_BG_ls	6	13.2%	13.8%	11.5%	20.9%	17.7%	7.5%	60.5%	41.1%
GE_Zoning_2_1_1_2_1_centroids 6-zone-1_BG_ls	6	13.6%	13.9%	12.2%	22.5%	17.4%	7.6%	60.5%	40.9%
GE_Zoning_1_2_2_2_1_centroids 7-zone-19_BG	7	13.7%	18.1%	13.4%	24.0%	12.5%	8.6%	64.5%	37.9%
GE_Zoning_2_1_2_2_2_centroids 7-zone-13_BG	7	13.7%	13.2%	13.9%	26.9%	15.4%	9.0%	60.5%	41.4%
GE_Zoning_1_2_2_2_1_centroids 7-zone-16_BG	7	13.7%	23.7%	8.9%	26.2%	12.4%	11.5%	64.8%	36.0%
GE_Zoning_2_1_2_2_2_centroids 7-zone-13_BG_ls	7	14.5%	14.4%	14.0%	26.1%	17.3%	9.0%	60.6%	41.9%
GE_Zoning_1_2_1_2_1_centroids 6-zone-1_BG	6	14.6%	23.4%	12.1%	26.5%	13.1%	10.4%	64.5%	37.0%
GE_Zoning_1_2_1_2_1_centroids 6-zone-9_BG	6	14.7%	14.9%	16.2%	24.3%	15.8%	9.5%	64.0%	39.0%
GE_Zoning_1_1_2_2_1_centroids 7-zone-15_BG	7	14.7%	19.1%	13.2%	23.8%	16.5%	8.7%	64.7%	41.3%

Fig. 16. The diversity selection stage.

the nuances of zone design. Users found the amount of new information overwhelming, and the researchers provided insufficient problem setup and too much information.

Despite these challenges, developing the zone selection tool provided valuable insights into how stakeholders engage with the zone design process and the appropriate level of technical complexity for effective engagement, and informed the prototype for the LLM-based interactive tool.

C ADDITIONAL RESULTS

C.1 Constraint Exploration

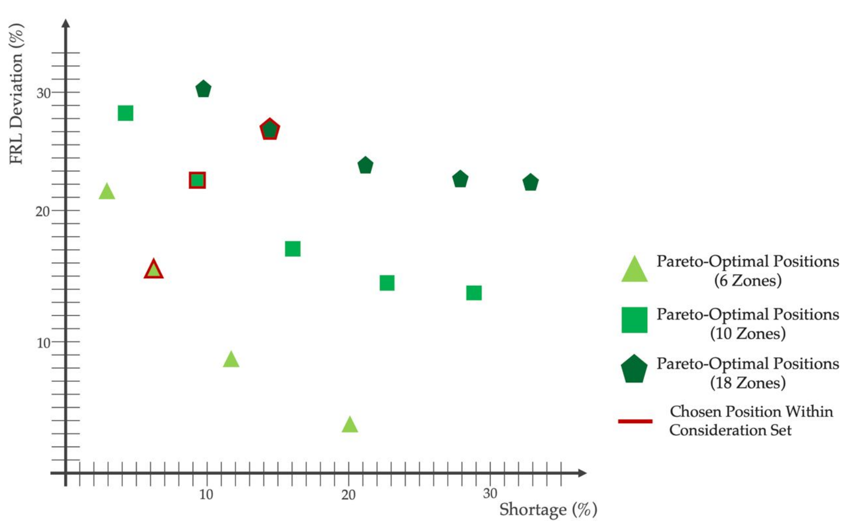


Fig. 18. Set of Pareto Optimal positions.

In Figure 18, we show the results of generating multiple solutions across different zone sizes to build a consideration set, systematically exploring trade-offs between Shortage (%) and FRL Deviation (%).

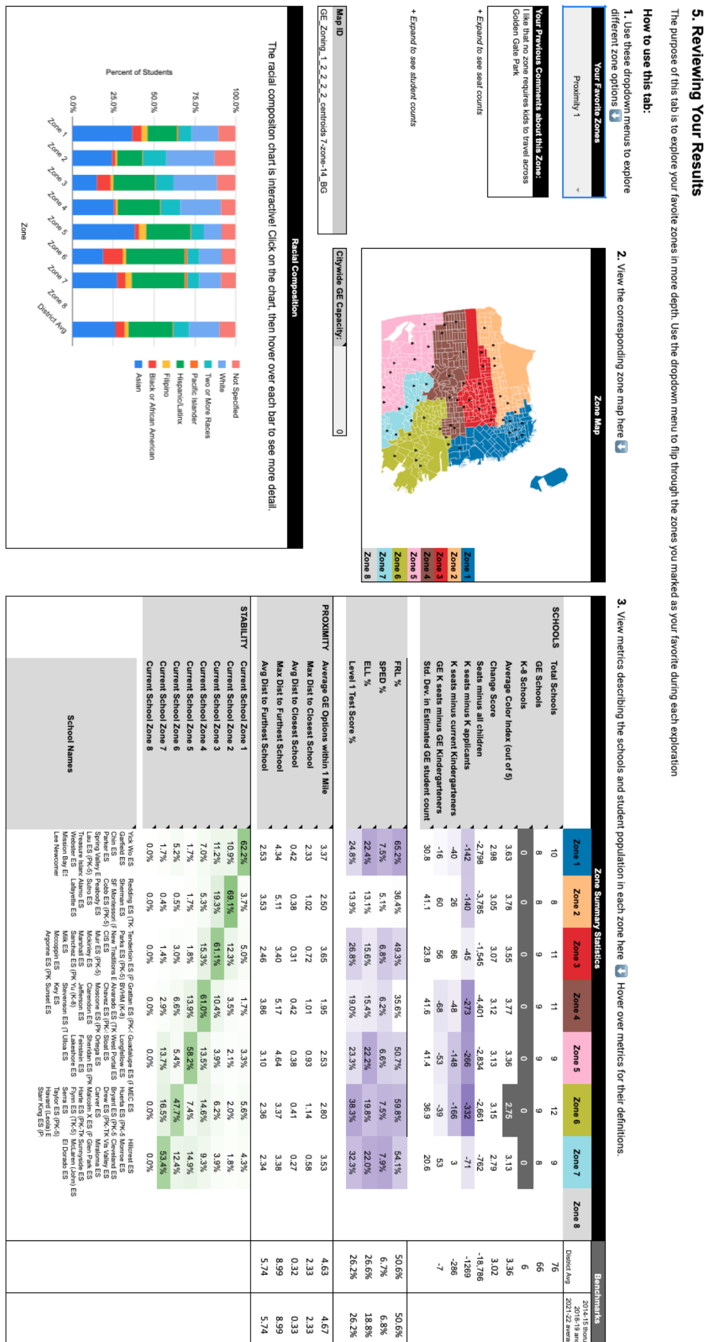


Fig. 17. The in depth exploration phase (partial) user interface. The relative performance of the current zone on diversity or proximity metrics compared to other zones in the family of solutions.

For each zone size, we enforced a series of maximum shortage thresholds—for example, ranging from 4% to 20% in our 6,10, and 18-zone solutions—while maintaining our hard constraints of contiguity and school balance. For each shortage threshold, we then optimized to find solutions with the lowest possible FRL Deviation while balancing other metrics of interest.

The resulting Pareto frontiers in Figure 18 demonstrate an inherent tension: solutions that achieve lower shortage percentages necessarily result in higher FRL deviations, illustrating that reducing capacity constraints comes at the cost of socioeconomic balance across zones. While the figure focuses on two key metrics for clarity, our full consideration set encompassed additional metrics and zone sizes.

C.2 Score-Based Filtering Approach

The scoring and filtering process relied on 10 working definitions: 5 diversity goals, 3 proximity goals, and 2 predictability goals.

(1) Diversity – Free or reduced price lunch dissimilarity
(2) Diversity – Racial dissimilarity
(3) Diversity – English language proficiency dissimilarity
(4) Diversity – Special education dissimilarity
(5) Diversity – Academic performance dissimilarity via level 1 (lowest) standardized test scores
(6) Proximity – Average walk time to all schools in zone
(7) Proximity – Percent of students within 20 minute walking distance of 2 or more schools
(8) Proximity – Percent of students within 20 minute transit distance of 2 or more schools
(9) Predictability – General education seat deficit
(10) Predictability – Total seat deficit

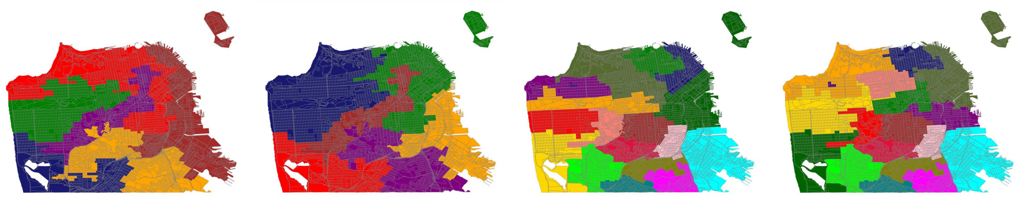
Each working definition was folded into a total map score via a weighted average. Each diversity objective was 10% of the total score (diversity giving 50% contribution), each proximity score was worth 6.67% of the total score (proximity giving a 20% contribution), and each of the the predictability goals was worth 15% of the total score (predictability contributing 30% of the score). While this scoring system was effective at narrowing the field of candidate maps, future research is needed to understand how to enable greater stakeholder participation and aggregate input from multiple stakeholders in the filtering process.

Number of Zones	FRL Balance	Racial Representativeness	Shortage	**Initial Count**	**Final Selection**
6	0.10	0.12	0.1	16	No
		0.15	0.1	18	No
	0.15	0.12	0.07	20	Yes
			0.1	14	No
		0.15	0.07	22	Yes
			0.1	20	No
7	0.10	0.15	0.1	8	No
	0.15	0.12	0.7	18	Yes
			0.1	12	No
		0.15	0.7	18	No
			0.1	18	Yes
10	0.15	0.15	0.15	2	No
		0.2	0.2	6	No
	0.22	0.15	0.1	3	No
			0.15	4	No
		0.20	0.1	11	Yes
			0.15	7	No
13	0.20	0.15	0.3	5	No
		0.2	0.2	4	No
	0.25	0.25	0.12	8	Yes
18	0.22	0.25	0.28	3	No
	0.26	0.25	0.14	5	Yes

Table 2. Results of Constraint Exploration and Constraint Selection. The column 'initial count' provides the number of positions in the consideration set, and the column 'final selection' indicates whether any of those positions were selected via the scoring process. At most one zone map was selected for each position.

C.3 Additional Relaxed ReCom Results

In this section, we provide for illustrative purposes some additional 6-zone and 18-zone maps generated by Relaxed ReCom (Figure 19), as well as a table showing the performance on balance constraints by one such zone map (Table 3). These maps exhibit many of the shortcomings discussed in Section 5.4.2 (e.g., non-compact zones with narrow bridges, variability in zone size, significant imbalance for multiple criteria).



(a) (6-Zones) (b) (6-Zones) (c) (18-Zones) (d) (18-Zones)

Fig. 19. Additional zone maps generated by Relaxed ReCom.

	Asian	**Black**	**Latinx**	**White**	**FRL%**	**Shortage**	**Number of Schools**
Zone 0	27.44%	6.58%	13.12%	28.03%	41.67%	16.22%	10
Zone 1	33.97%	1.43%	10.20%	30.02%	29.88%	17.98%	9
Zone 2	30.81%	2.55%	22.46%	21.88%	44.63%	-36.75%	11
Zone 3	13.21%	19.21%	24.53%	20.23%	59.35%	10.03%	9
Zone 4	24.64%	2.94%	18.80%	31.11%	42.95%	-39.10%	12
Zone 5	37.37%	5.77%	25.97%	10.52%	59.57%	19.70%	9
Max Deviation	16.18%	14.24%	8.50%	13.53%			

Table 3. Performance on metrics of interest (e.g. demographic distribution, shortage per zone) of one well-performing 6-zone map obtained using Relaxed ReCom. Max difference in school counts across zones is 3. Maximum shortage is 19.7%, and maximum FRL deviation compared to district average is 15.03%.

C.4 Additional Metrics and Results for Choice-Driven Evaluation

In this section, we provide results for additional metrics in choice-driven evaluation of zone-based policies. Additional proximity metrics include the percentage of students assigned further than 3 miles from home. Additional diversity metrics include the dissimilarity index for high-FRL students, which is the proportion of students living in high-FRL blocks (those with $\geq$ 50% of student residents eligible for free- or reduced-price lunch) who would need to be reassigned for the assignment to be representative. Additional choice metrics include the proportion of students who were not assigned to one of their choices (i.e., 'designated' to a program not on their list or not assigned at all), assigned their top choice. assigned one of their top 3 choices within their zone, and assigned a 'bad' choice ($\geq$ 3 miles from home and ranked 5th or worse on their preference list). We also included predictability metrics, such as variability in assigned assigned program rank and distance, and the percentage of students without CTIP1 priority who are assigned outside their top 3 choices (this reflects assignment uncertainty for students lacking priority). (CTIP1 priority gives priority at all schools to students who live in census blocks with historically low test scores. These are determined using Level 1 standardized test score rates from the 2010 census.)

		6 Zones + Reserves	6 Zones + Reserves v2	10 Zones + Reserves	13 Zones + Reserves	18 Zones + Reserves	New AAs	Status Quo
Proximity	*Avg Distance (miles)*	1.31	1.32	1.20	1.18	1.19	1.33	1.31
	Distance ≤ 0.5 miles	31.2%	32.2%	35.4%	35.7%	36.5%	34.9%	36.0%
	Distance ≥ 3 miles	9.6%	10.4%	8.5%	8.2%	8.6%	11.1%	10.9%
Diversity	*+15% FRL Schools*	11.9	12.2	14.9	14.2	14.7	13.3	12.6
	% AALPI in +15% FRL	15.7%	16.0%	19.9%	21.3%	22.3%	19.8%	18.5%
	Dissimilarity (High FRL)	21.3%	20.9%	23.1%	22.8%	23.6%	20.2%	20.4%
Choice	*Designated or Unassigned*	6.74%	6.64%	6.70%	6.88%	7.82%	5.57%	5.50%
	Prop Top 1 choice	58.0%	58.9%	57.4%	54.6%	52.8%	65.6%	65.6%
	Prop Top 3 choice	80.4%	81.2%	80.2%	77.1%	74.9%	87.7%	87.9%
	Prop Top 3 choice (In Zone)	88.5%	89.6%	90.5%	89.0%	89.1%	87.7%	87.9%
	Dist ≥ 3, Rank ≥ 5	1.43%	1.77%	1.02%	0.92%	1.22%	1.26%	1.22%
Predictability	*Variance of In-Zone Rank (All Assigned)*	6.69	6.05	3.92	4.97	4.75	2.10	2.06
	Variance of Distance (All Assigned)	1.34	1.41	1.28	1.25	1.30	1.50	1.50
	Rank Top 3 (In Zone, non-CTIP)	86.4%	87.7%	88.8%	87.0%	87.0%	85.5%	85.7%

Table 4. Average assignment metrics using preferences generated by choice model. In each row, all policies improving on the Status Quo for a given metric are colored in blue.

These results confirm the main insights from Table 1 in the body of the paper. Zone-based policies, especially with 10 or fewer zones, generally improve distance metrics compared to the status quo. Only the 6-zone policies noticeably improve on diversity metrics compared to the status quo. All zone-based policies reduce choice metrics that consider out-of-zone preferences, with the caveat that the students at the tails assigned ‘bad’ choices decreases with mid-sized zones (10-13 zones). In addition, introducing zones reduces variability of distance and in-zone rank of assigned programs.