Reducing Inter-District Inequality as a Constrained Multi-Objective Optimization Problem

Gabriel H. Bermudez, Hyun-Jeong Lee, Conor E. Olive, Kyle Wright[1], Cory S. McCullough[1]

[1] Graduate students

University of California, Merced

Introduction

- Prior to 1970, there was an increasing trend in school districts consolidation
- After 1970, communities within consolidated school districts increasingly seceded to form their own more homogeneous districts, with 41 successful examples (Kitchens 2021)
- How do we find optimal ways to decrease inter-district inequalities for multiple variables (such as segregation)?

Model

- We consider the movement of students as x, where x is a square matrix
- Each element $x_{i,j}$ denotes the percentage of students moved from district i to district j
- We model our problem as the constrained, nonlinear, multi-objective optimization problem in (1)

$$\begin{aligned} \min_{x} \quad & \frac{1}{N}\sum_{n=1}^{N}\left[f_\alpha(x)+f_\beta(x)+f_\gamma(x)\right] \\ \text{subject to} \quad & c_1(x) \le p \\ & c_2(x) = 1 \\ & 0 \le x_{ij} \le b_{ij}, \quad i,j = 1,\dots,N \end{aligned} \tag{1}$$

- where N is the number of districts
- where $c_1(x)$ gives the percentage of students moved and p upper bound on this percentage
- where $c_2(x)$ gives a $N \times 1$ column vector where each row is the sum of percentage of students leaving district i and staying in district i
- setting $c_2(x) = 1$ ensures conservation of students

$$b_{ij} = \begin{cases} 1 & \text{districts } i \text{ and } j \text{ share borders} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

- f_α, f_β, and f_γ measure mean square deviation (MSD) for race, student to teacher ratio, and revenue from local sources per student, respectively
- Each function is normalized using min-max scaling

$$f_\alpha(x) = \sum_{r=1}^{R}\left(\alpha_{r,n}(x) - \mu_r\right)^2 \tag{3}$$

- R is the number of races
- $\alpha_{r,n}$ gives the new percentage of race r in district n
- μ_r gives the percentage of students of race r across all districts

$$f_\beta(x) = \left(\beta_n(x) - \mu_\beta\right)^2 \tag{4}$$
$$f_\gamma(x) = \left(\gamma_n(x) - \mu_\gamma\right)^2 \tag{5}$$

- β_n gives the new student-to-teacher ratio in district n
- μ_β gives the mean student-to-teacher ratio across all districts

Model (cont.)

- γ_n gives the new local revenue per student in district n
- μ_γ gives the mean revenue from local sources across all districts
- We use a sequential quadratic programming method, implemented in `scipy.optimize.minimize(method=SLSQP)`, to minimize our objective function (Virtanen et al. 2020)
- Validation of the method on our model is shown in Figure 1

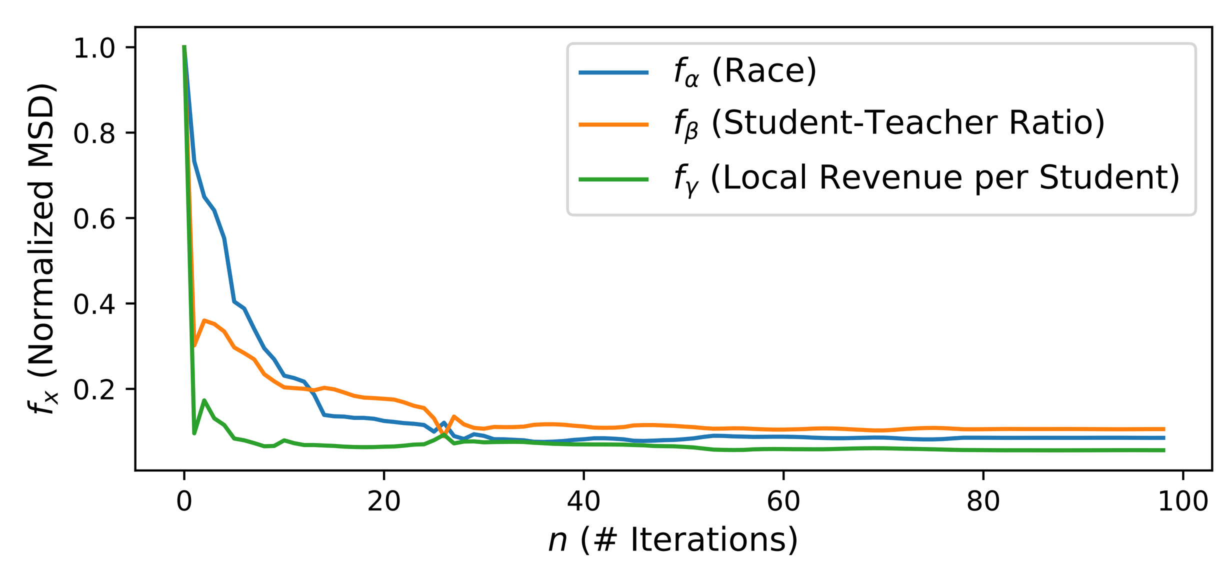

Fig. 1: Change in $f_\alpha, f_\beta, f_\gamma$ with increasing iterations

Results

- We use 2018-19 data from unified school districts in Alameda County as a test case (*Common Core of Data, Education Data Portal (Version 0.18.0)* 2023)
- Figure 2 shows the relative reduction in inequality between districts as p is increased
- Increases in constraint p provides diminishing returns, with full minimization occurring around 0.7

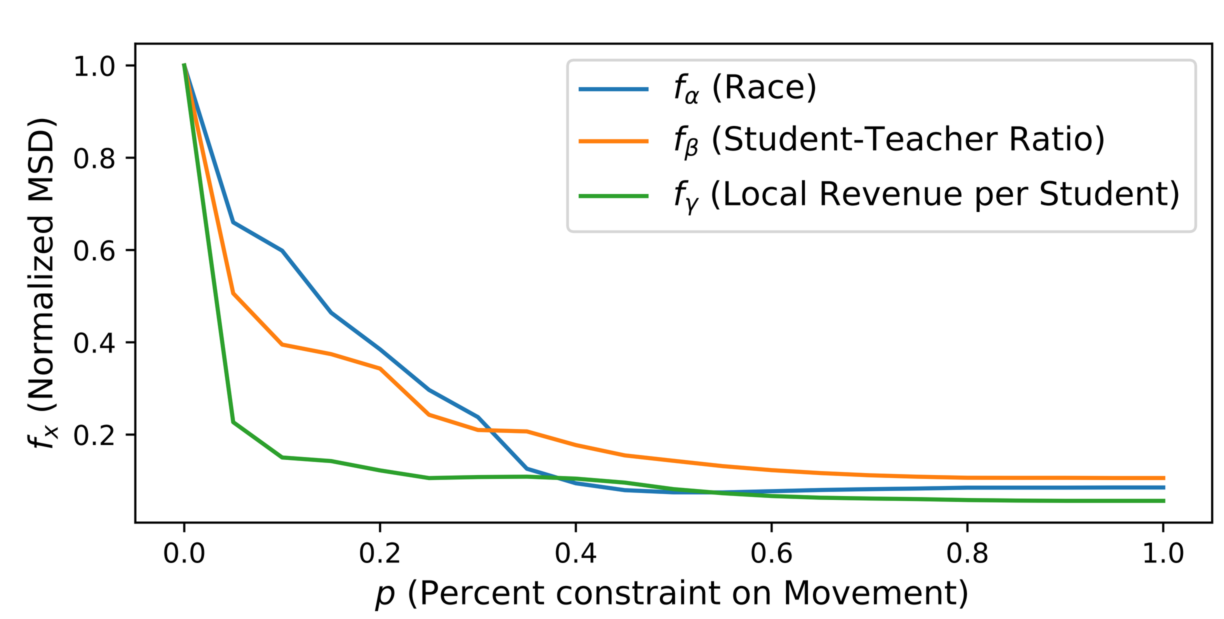

Fig. 2: Normalized MSD for different district-level variables at various constraints p

Results (Cont.)

- Figure 3 and 4 show the normalized change in f_α when $p = 1$
- With $p = 1$ maximum race MSD is decreased by 73.7%
- Mean race MSD also is decreased by 73.1%

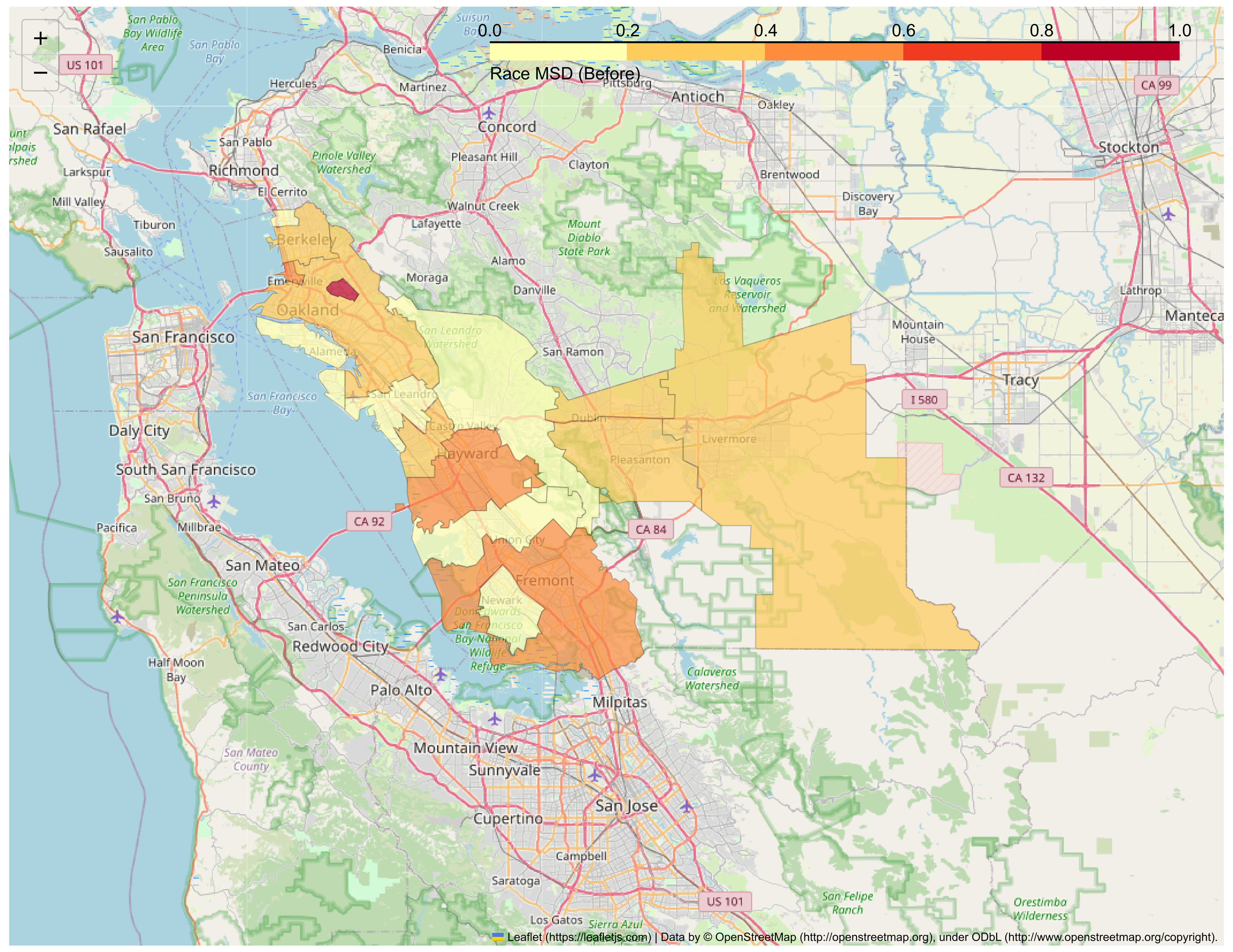

Fig. 3: Race MSD for Alameda County pre-optimization

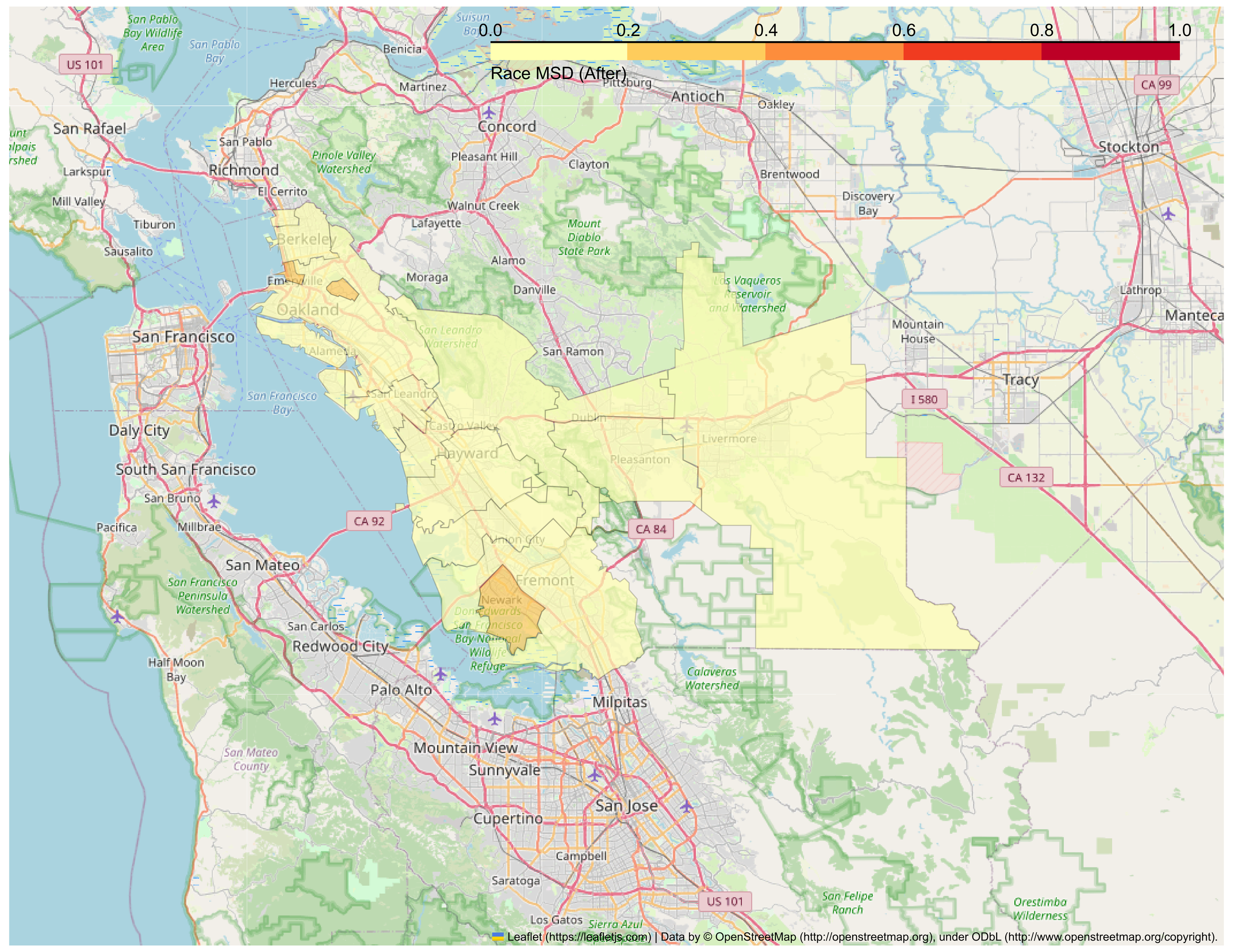

Fig. 4: Race MSD for Alameda County post-optimization

Discussion

- Though Alameda county is chosen as a test case, this model can be applied to any general region
- Since each objective function is simply calculating the MSD, additional variables may be optimized in a trivial manner
- Future work may be done to automatically generate new district boundaries that facilitate the movement of students to new districts
- Future work may be done to discretize geographical units within districts where within-district populations are geographically non homogeneous

References

Common Core of Data, Education Data Portal (Version 0.18.0) (2023). URL: https://educationdata.urban.org/documentation/.

Kitchens, Karin (2021). "Dividing Lines: The Role School District Boundaries Play in Spending Inequality for Public Education". In: *Social Science Quarterly* 102.1, pp. 468–491. DOI: https://doi.org/10.1111/ssqu.12886.

Virtanen, Pauli et al. (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17, pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. (DMS-1840265)