Community- and Data-Driven Homelessness Prevention and Service Delivery: Optimizing for Equity

Amanda R. Kube, Ph.D.

Data Science Institute, The University of Chicago;

Division of Data and Computational Sciences

Washington University in St. Louis, MO

Sanmay Das, Ph.D.

Department of Computer Science

George Mason University

Washington DC

*Patrick J. Fowler, Ph.D.

Brown School of Social Work, Public Health, and Social Policy;

Division of Data and Computational Sciences

Washington University in St. Louis, MO

*Corresponding Author
Patrick J. Fowler, Ph.D.
One Bookings Drive
MSC 1196-251-46
St. Louis, MO 63130
email: pjfowler@wustl.edu
Office 314-935-5859

Abstract

Background. To test data-driven and equity-focused homelessness prevention. Federal policies call for efficient and equitable local responses to homelessness. However, the overwhelming demand for limited housing assistance challenges efforts to prioritize services, and little evidence supports decision-making. We demonstrate a community- and data-driven approach for prioritizing scarce resources.

Methods. Administrative records captured homeless service delivery and outcomes in St. Louis, MO from 2009-2014 (n=10043). Counterfactual machine learning identified services most likely to prevent household-level homelessness within two years, which we aggregated to design group-based service prioritization rules. Simulations re-allocated households to available services and evaluated whether data-driven prioritization reduced community-wide homelessness without excluding marginalized and underrepresented groups.

Findings. Local data showed households with comorbid health conditions avoided homelessness most when provided longer-term supportive housing, and families with children fared best in short-term rentals. Prioritization rules reduced community-wide homelessness and disproportionately benefited marginalized and minoritized populations.

Interpretation. Leveraging local records with machine learning supplements local decision-making and enables ongoing evaluation of data- and equity-driven homeless services. Community- and data-driven prioritization rules more equitably target scarce homeless resources.

Funding. The work was supported by the National Science Foundation through awards 2127752, 2127754, and 1939677, and from Amazon through an NSF-Amazon Fairness in AI award. Sponsors had no role in the study; collection, analysis, and interpretation of data; writing the report; and the decision to submit the report for publication.

Growing housing insecurity and homelessness rates compromise public health – disproportionately burdening underrepresented and marginalized populations. The overwhelming demand for scarce housing resources compromises federal policies requiring communities to develop equitable and cost-efficient responses to homelessness. Local service providers continually face difficult decisions on who to serve first. Although HUD guidelines emphasize vulnerability, serving the worst-off first, the realities of severe resource constraints and inaccurate and racially biased risk assessments compromise decision-making [1–4]. A recent AJPH Opinion by Shinn and Richard (2022) identifies the inherent ethical tradeoffs of prioritizing one group over another for scarce services (e.g., serving homeless infants before homeless elderly), empirically demonstrated in other work [5–8]. As one example, policies that set preferences for homeless infants before homeless elderly require communities to engage in values-driven system design that exceeds the data-driven capacities of coordinated assessment strategies.

Questions remain regarding how best to design local homeless services that achieve equity and efficiency, as noted in Shinn and Richards (2022). A progressive engagement approach recommends initially offering households the least intensive service (e.g., housing case management) and leveling up to more intensive interventions (e.g., supportive housing) until achieving stability; however, progressive engagement generates inefficiencies from service mismatch that burden households seeking needed supports. Another two-staged approach assesses household risk that feeds decision tools for customizing services tailored to meet household needs[10]. Algorithmically driven approaches also show promise for targeting homeless services. Using linked administrative records, Alleghany County, Pennsylvania, prioritizes housing assistance for households predicted to be booked in jail and need emergency healthcare that benefits Black households more than standard vulnerability assessments[11]. Research leveraging machine learning with community-wide Homeless Management Information System (HMIS) data illustrates the potential for identifying and matching households to the available intervention most likely to prevent future homelessness[8].

The present study extends recent innovations on data-driven homeless system design to integrate advances in equity-based artificial intelligence[12–14]. Based on precision public health[15-18], counterfactual machine learning identifies services most likely to benefit similar individuals by examining heterogeneous treatment effects from observational data[19]. Community-wide administrative records capture an array of homeless services delivered with outcomes over time that feed algorithmic predictions of household-level response to interventions – moving further upstream than predicting risk in the absence of intervention[8, 20]. Although the aforementioned resource constraints prohibit fully automated service allocations, aggregating household-level predictions for subgroups of interest provides communities with additional evidence for decision-making. Any allocation of scarce resources inherently produces winners and losers[5, 6, 21], but services guided by transparent prioritization rules allow ongoing assessment of pre-determined thresholds for equity and efficiency tradeoffs[6, 21]. We offer a proof-of-concept case study based on long-standing work within St. Louis, MO and consider central issues for local implementation.

Methods

A homelessness management information system (HMIS) tracked all requests and delivery of homelessness services from 2009 to 2014 across St. Louis, MO – a medium-sized Midwestern legacy city facing historical segregation and ongoing socioeconomic challenges [22-24]. The present study includes first-time entries into four available services: 1) homelessness prevention provides one-time financial assistance for families at imminent risk of homelessness; 2) emergency shelter offers time-limited group accommodations to avoid staying on the streets; 3) rapid rehousing – initially implemented in October 2009 – offers homeless households short-term community-based rentals for up to 12 months; and 4) transitional housing gives up to two years of congregate accommodations with social services. During the study period, first-time enterers did not receive permanent supportive housing, a service that offers longer-term and more intensive support.

The longitudinal data also captures reinvolvement with homeless services over time across the more than 75 community-based programs serving the homeless in St. Louis. Data represent homeless services before federally required coordinated assessment and entry initiatives; generally, households received services on a first-come-first-serve basis without household vulnerability assessment. Moreover, St. Louis used a central hotline to coordinate all requests for housing assistance recorded in the HMIS, and thus, the data captured community-wide demand for services unconstrained by availability.

Analyses predict whether households reentered homeless services – either requesting or receiving additional homeless services within two years of initial exit from the system – coded as a binary outcome. Predictors include 35 household characteristics collected upon initial entry into the services. Data record service requests, entry and exit dates, and an array of household sociodemographics (e.g., age, gender, race/ethnicity, income), functioning (e.g., chronic conditions, mental illness, substance abuse), and household characteristics (e.g., number of children, adults, prior living arrangements). Features come from HUD and federally required HMIS universal and program-specific data elements. Appendix A describes all features with summary statistics by initial homeless service received.

We develop and test the design of equity- and data-driven homeless systems in four phases. First, counterfactual predictions generate household-level probabilistic estimates of system reentry given receipt of homelessness prevention, emergency shelter, rapid rehousing, and transitional housing. We use Bayesian Additive Regression Trees (BART) – a machine learning approach that demonstrates bias reductions in observational and complex data and outperforms propensity score and nearest-neighbor matching algorithms [19, 25]. BART outputs 1000 sample estimates of the probability of reentry for each household, and average treatment effects (ATE) represent the differences in reentry probabilities between every pairwise combination of services (e.g., shelter vs. rapid rehousing, rapid rehousing vs. transitional housing). Estimates include the mean and 2·5% and 97·5% quantiles for a 95% estimated credible interval of the pairwise differences. We restrict comparisons of homelessness prevention to households who did not meet federal criteria for homelessness upon initial entry into services; this addresses a potential confound in a counterfactual that not only provides prevention but also conceptually takes homes away from these households.

Second, we generate homeless service prioritization rules through iterative exchanges with community partners. Initial hypotheses articulate the subpopulations most likely to do best in available services. No household is expected to do best in emergency shelter, although some could be especially harmed. Households with comorbid conditions – operationalized as self-reporting at least two disabilities, mental health, and substance abuse problems – should be prioritized for transitional housing that attends to psychosocial barriers to stability, whereas families with children under 18 years of age are more likely to face socioeconomic drivers of instability better addressed with less intensive rapid rehousing. Community partners also consider service effectiveness for several subpopulations without clear hypotheses. Unaccompanied youth aged 18-to-24 years without children vary in need with some youth appearing to do best with transitional housing and others in rapid rehousing depending on their circumstances. Likewise, concerns for minoritized and marginalized populations warrant comparisons of Black versus Whites, and females versus males. For subpopulations of interest, we estimate conditional average treatment effects (CATE) that aggregate household-level ATE of transitional housing versus rapid rehousing and provide group-level means and 95% credible intervals. A review of CATE across subpopulations of interest informs the formulation of prioritization rules that preference groups who do better in transitional housing or rapid rehousing.

Third, we evaluate the efficiency of prioritization rules by simulating optimal homeless service delivery under a set of conditions. Simulations using an integer programming framework consider household-level service effectiveness and dynamic resource constraints limiting matching with services that most reduce homelessness. Specifically, we use BART out-of-sample system reentry predictions given receipt of each service plus capacity limits – derived from multiplying the actual number of households who entered each service type in a week by the average weekly costs of each service derived from prior research and adjusted to 2022 inflation[26-29]. Appendix B reports the weekly average costs and the calculation of the overall capacity limit. A weighted bipartite matching algorithm assigns households entering services each week to one of the four services that minimize homeless reentries. Prioritization rules simply serve households from preferred subpopulations first until the exhaustion of resources. We assess prioritization allocations on system-wide efficiency (reentry reductions) and cost-effectiveness (total service expenditures) compared with 1) the original allocation (services-as-usual), 2) optimizing for minimal costs, and 3) optimizing for reducing reentry.

Finally, we evaluate the equity of the community- and data-driven prioritization rules. We operationalize equity using two commonly referenced metrics for scarce resource allocation – the utility gains and shortfalls experienced by subpopulations of interest [6, 21, 29]. Utility represents the predicted probability of homelessness reentry under each potential service allocation. Gain compares the utility with a worst-case scenario of serving all households with the service predicted to prevent homelessness the least, and thus, measures the benefit gained from each allocation. Shortfall contrasts with the best-case scenario when all households receive the service most likely to prevent homelessness – assessing the loss generated by each allocation. We assess change in gain (ΔG) and change in shortfall (ΔS) for allocation **a** compared with allocation **r** which randomly allocates households to services; the differences are calculated for subpopulations of interest (s = 1) versus a comparison group (s = 0) as follows:

$$\Delta G(\mathbf{a}) = (E[\frac{1}{N_0}\sum_{i,s=0} \mathbf{a}.\frac{u_i}{u_i^{min}}] - E[\frac{1}{N_1}\sum_{i,s=1} \mathbf{a}.\frac{u_i}{u_i^{min}}]) - (E[\frac{1}{N_0}\sum_{i,s=0} \mathbf{r}.\frac{u_i}{u_i^{min}}] - E[\frac{1}{N_1}\sum_{i,s=1} \mathbf{r}.\frac{u_i}{u_i^{min}}])$$

$$\Delta S(\mathbf{a}) = (E[\frac{1}{N_0}\sum_{i,s=0} \mathbf{a}.\frac{u_i}{u_i^{max}}] - E[\frac{1}{N_1}\sum_{i,s=1} \mathbf{a}.\frac{u_i}{u_i^{max}}]) - (E[\frac{1}{N_0}\sum_{i,s=0} \mathbf{r}.\frac{u_i}{u_i^{max}}] - E[\frac{1}{N_1}\sum_{i,s=1} \mathbf{r}.\frac{u_i}{u_i^{max}}])$$

Random allocations represent the average results across 100 runs with all all prevention eligible households receiving prevention for consistency. The difference in gain and difference in shortfall equal zero when no bias exists across groups. By centering the fairness metrics, equity values of 0 indicate the allocation has equal bias to that of a random allocation. In our calculations, negative numbers indicate greater gain and shortfall for subpopulations of interest, while positive numbers signal comparison group preferences. We expect equitable homeless services to demonstrate greater gains and greater shortfalls for subpopulations of interest.
We hypothesize the following:

- Prioritization rule-based allocations increase access to appropriate services for preferred groups without increasing overall system reentries and service costs.
- Prioritization achieves equitable access to appropriate services for traditionally underrepresented and marginalized subgroups

Results

Administrative records track 10043 households who initially entered homelessness prevention (49·3%) or homeless services (50·7%) between 2009 to 2012 in St. Louis. Half of the households (49·3%) are eligible for homelessness prevention given housing insecurity, while others require emergency housing. The ineligible population disproportionately includes Black (85·0%) and female-headed (66·6%) households aged 39·5 years on average (SD=12·8) with 10·7% unaccompanied homeless youth aged 18 to 24 years. Families with children comprise 58·3% of households. Nearly one-fifth of household heads self-reported a disabling health condition (15·0%), mental health problem (29·8%), or substance abuse (27·4%) upon entry into services.

Overall, 27·5% of households request or reenter services. Households initially referred to homelessness prevention (n = 4737) and transitional housing (n = 1469) reenter less (13·6% and

34·3%, respectively) compared with rates for households in emergency shelter (42·5% of 2997 households) and rapid rehousing (40·6% of 840 households). As is common in observational data, pre-existing differences at the time of service entry, as seen in Appendix A, contribute to differences in reentry as well. Bayesian additive regression trees (BART) have been shown to mitigate this and so were used to model complex interactions and nonlinearities in the data and generate counterfactual pairwise reentry predictions. The model demonstrates adequate fit and accuracy when predicting reentry with an Area Under the Receiver-Operating-Characteristic Curve (AUC) of 0·75, misclassification error rate of 0·2, precision of 0·6, recall of 0·3, and calibration of 0·9 – all metrics suggest acceptable accuracy in the prediction task. Furthermore, the BART-generated out-of-sample counterfactual predictions of reentry given the actual service provided correspond closely with the observed reentry rates across interventions.

We initially examine population average treatment effects (ATE) to inform prioritization rules. Counterfactual predictions suggest all 4942 homelessness prevention-eligible households do “best” (lowest probability of reentry) in prevention with pairwise ATE showing a 5·54 percentage point (pp) reduction in reentry compared with transitional housing, 6·04 pp reduction for rapid rehousing, and 7·80 pp reduction for emergency shelter. For the 5101 prevention-ineligible households, treatment effects vary more; 65% are predicted to do best in transitional housing, 30·1% in rapid rehousing, and 4·2% in emergency shelter. The treatment effect of transitional housing compared with rapid rehousing on homelessness reentry fell close to zero with considerable variation (ATE = -0·02, 95% Credible Interval = -0·12, 0·07). Together, evidence from population effects supports prioritizing all eligible households for prevention and suggests examining potential subpopulation differences for prevention-ineligible households.

Figure 1 illustrates the conditional average treatment effects (CATE) of transitional housing compared to rapid rehousing for homeless subpopulations identified by community stakeholders. Treatment effects that fall further from the ATE represented by the dashed line indicate better response to rapid rehousing (above) and transitional housing (below). No subpopulation does best in shelter, and thus, we ignore those pairwise estimates. Surprisingly, households with mental health problems do relatively worse in transitional housing, and thus, we reoperationalize comorbidities as any two of alcohol, drug, or disability. CATE support community hypotheses that households with comorbid conditions exhibit lower reentry rates given transitional housing compared with rapid rehousing, averaging 6·1 pp reduction. Results also show that families with children and without comorbidities exhibit a 3·2 pp reduction in reentry when receiving rapid rehousing, which supports the community hypothesis. The CATE for other subpopulations and intersectional identifies fall close to the average treatment effect indicating no clear service preference, as presented in Appendix C. Generally, results show households with comorbidities do better in transitional housing across identities, while households with children and without comorbidities do better in rapid rehousing.

Figure 2 visualizes the resulting community- and data-driven prioritization rules. Homelessness prevention-eligible households receive prevention. Non-prevention-eligible households with comorbid health, mental health, or substance abuse problems receive transitional housing if accessible in the week the household enters the system. Families with children under 18 years of age without comorbidities receive rapid rehousing if available. Other non-prevention-eligible households enter a lottery for emergency shelter, rapid rehousing, or transitional housing. The lottery assigns as many households as possible receive transitional housing as the ATE for transitional housing versus rapid rehousing indicates the general population does better in transitional housing, while maintaining the same cost of service provision as in the services-as-usual allocation.

Table 1 reports simulation results assessing the implementation of prioritization rules compared with services-as-usual, random assignment, service-efficiency optimization, and cost-effectiveness optimization. Prioritization rules reduce system reentries by one pp compared with services-as-usual, and the results correspond with cost-effectiveness optimization and yield similar budget savings compared with services-as-usual. The expected small efficiency gains from prioritization and cost-effectiveness optimization reflect budget constraints and ongoing demand that requires all households to receive assistance. Optimizing on service efficiency lowers reentries by 2·2 pp compared with services-as-usual – helping 113 households avoid homelessness during the follow-up; however, reductions require over $1 million, a 25·1% increase in total budget. Prioritization results suggest minimal compromise on system efficiency.

Figure 3 presents the equity of community- and data-driven prioritization rules. Bars plot the difference in gains (left) and shortfall (right) for subpopulations of interest versus majority groups (listed vertically). Results show that prioritization rules yield the largest gains and shortfalls for households with comorbidities compared with cost-effectiveness and services-as-usual. Thus, prioritization disproportionately favors households with comorbidities when giving their most and least useful service, which enhances equity. For families with children but without comorbidities, the utility gained from prioritization is similar to services-as-usual that favored families, however, allocating services based on efficiency disproportionately burdens families, making services less equitable. We also present equity metrics by race/ethnicity, gender, and unaccompanied youth. None of the allocation schemes disproportionately favor Black households, which represent four of every five households. Female-headed households are disproportionately burdened by all allocations; however, the burden is substantially mitigated by prioritization rules compared with services-as-usual and cost efficient assignments. Similarly, the burden on unaccompanied youth is mitigated by both prioritization rules and cost efficient assignments compared with services-as-usual. Overall, prioritization rules generally produce the most equitable homeless services – as intended.

Discussion

The study demonstrates the feasibility of an iterative community- and data-driven approach for effective and equitable homeless service delivery. Findings address an important gap between policy and practice [3, 8, 20]. Federal guidelines require communities to allocate scarce housing assistance based on system-wide assessments of household risk [1]; yet, little evidence supports the accuracy and cultural validity of existing tools [4, 5]. Furthermore, the scarcity of homeless services inherently requires homeless providers to make continual moral preferences on whom to serve first with little ability to evaluate individual decision-making and system goals[5].

Our approach elicits feedback from key stakeholders to define subpopulations of interest and relevant intersectional identities. In this pilot, target households initially included those with comorbid conditions, families with children, unaccompanied youth, African American, and female-headed. Leveraging historical administrative records, counterfactual machine learning shows transitional housing reduces reentries the most for households with comorbidities, and families with children and no comorbidities do best with rapid rehousing during the study period, regardless of race/ethnicity, gender, and age. The evidence informs transparent, easily implementable, and evaluable prioritization rules for targeting services that minimize system reentries. Simulations of prioritization rules demonstrate larger reductions for subpopulations of interest (i.e., comorbidities and families) without perpetuating disparities by race/ethnicity, gender, age, and unaccompanied youth. Results demonstrate promise for homeless service delivery that incorporates community- and data-driven insights.

Findings must be interpreted in the context of several conceptual and methodological limitations. First, targeting fails to address the lack of affordable housing that drives the overwhelming demand for housing assistance; the problem of whom to serve first remains unsolved. Although we demonstrate an approach for articulating and evaluating ethical preferences for scarce resource allocation, reforms that make safe and affordable housing accessible to low-income households remain critical for just homeless service delivery. Second, we demonstrate feasibility of the approach and ignore the formidable implementation challenges associated with community- and data-driven homeless services. For instance, our simulations automate resource allocation in ways that would and should not be incorporated into service delivery. Homeless consumers and providers access relevant information for decision-making that fails to appear in HMIS, and data-driven services should elicit and incorporate preferences by directly adjusting algorithms and through designing mechanisms that allow adjustments and overrides that ultimately improve tailored resource allocation [30-32]. Moreover, the introduction of data-driven decision supports into homeless services introduces non-trivial dynamics on how information is interpreted and used that could generate unexpected outcomes [21, 33]. Rigorous research must consider the intended and unintended consequences of prioritizing scarce resources.

Finally, a series of technical issues limit insights from the modeling. Our data capture households during recovery from the Great Recession and fail to account for the current social and service delivery contexts. Noteworthy, the data predate federal initiatives around coordinated entry and housing first, and instead, allocations functioned primarily as first-come-first-serve. Generating unconfounded treatment effects could prove difficult with current data that includes changing system preferences. Likewise, federal system performance measures now focus on service receipt and not need, whereas our outcome considers all re-requests for assistance regardless of availability that might be less prone to systematic exclusion from services. Lastly, model building requires considerable local tailoring that meet community interests and the statistical assumptions necessary for counterfactual estimation, such as measurement quality, sample size, statistical power, etc. HMIS collects selected household features that might not capture the highly dimensional mechanisms underlying service delivery and intersectionalities of interest. The iterative approach requires deep collaboration on the technical and substantive elements of model building.

In sum, the study demonstrates the feasibility for community- and data-driven homeless service delivery that maximizes resources with explicit attention to disparities around minoritization and marginalization. The encouraging results require continued development of technical and ethical capacities for implementation. Accessible affordable housing remains a fundamental issue for promoting housing security for low-income households.

Authors' Contributions

Amanda Kube: Conceptualization, Data Curation, Formal analysis, Methodology, Validation, Visualization, Writing – Original Draft, Writing – Review and Editing
Sanmay Das: Conceptualization, Data Curation, Formal Analysis, Methodology, Validation, Visualization, Writing – Original Draft, Writing – Review and Editing, Supervision, Funding Acquisition
Patrick Fowler: Conceptualization, Data Curation, Formal analysis, Methodology, Validation, Visualization, Writing – Original Draft, Writing – Review and Editing, Supervision, Funding acquisition, Project Administration, Resources

Authors have directly accessed and verified the underlying data reported in the manuscript.

Data Sharing

Data generated for this study will be provided as a repository on GitHub. A CSV file consists of 13940 rows and seven columns. Each row represents a de-identified and homeless household. Columns list the factual and estimated counterfactual allocations for each service (ES, TH, RRH, and Prev) calculated by BART, the outcome of whether the household reentered homeless services within two years, whether the household was prevention eligible.

The columns labeled ES, TH, RRH, and Prev give the counterfactual prediction for whether that household would need services again within 2 years if they had been allocated to that intervention. ES stands for Emergency Shelter, TH Transitional Housing, RRH Rapid Rehousing, and Prev stands for Homelessness Prevention. Couterfactual predictions are calculated using BART, as described in detail in the paper referenced above.

The variable named PrevEligible is an indicator of whether (1) or not (0) a household is eligible to receive prevention services. This indicator is calculated using information about a household's previous residence and current housing situation. If a household is not eligiable to recieve prevention, a prediction for probablity of reentry if that household is placed in prevention is not provided. Use of the anonymous data should cite the present article

References

1. Department of Housing and Urban Development. Coordinated entry core elements. *HUD Exchange.* 2018. https://files.hudexchange.info/resources/documents/Coordinated-Entry-Core-Elements.pdf.

2. Gubits D, Shinn M, Wood M, Brown SR, Dastrup SR, Bell SH. What interventions work best for families who experience homelessness: Impact estimates from the family options study. *J Policy Anal Manag*. 2018;37(4):835-866. doi:10.1002/pam.22071

3. Shinn M, Brown SR, Spellman BE, Wood M, Gubits D, Khadduri J. Mismatch between homeless families and the homelessness service system. *Cityscape*. 2017;19(3):293-307. PMID: 29326756; PMCID: PMC5760190.

4. Brown M, Cummings C, Lyons J, Carrión A, Watson DP. Reliability and validity of the Vulnerability Index-Service Prioritization Decision Assistance Tool (VI-SPDAT) in real-world implementation. *J Soc Distress Homel*. 2018;27(2):1-8. doi:10.1080/10530789.2018.1482991

5.Shinn M, Richard MK. Allocating homeless services after the withdrawal of the Vulnerability Index–Service Prioritization Decision Assistance Tool. *Am J Public Health*. 2022;112(3):378-382. doi:10.2105/ajph.2021.306628

6. Mashiat T, Gitiaux X, Rangwala H, Fowler PJ, Das S. Trade-offs between group fairness metrics in societal resource allocation. *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency.* 2022; 1095-1105

7. Estornell A, Das S, Liu Y, Vorobeychik Y. Unfairness despite awareness: Group-fair classification with strategic agents. *arXiv preprint arXiv: 2112 02746.* 2021.

8. Kube A, Das S, Fowler PJ. Allocating interventions based on predicted outcomes: A case study on homelessness services. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2019;33(1):622-629.

9. Culhane DP, Metraux S, Byrne T. A prevention-centered approach to homelessness assistance: a paradigm shift? *Housing Policy Debate*. 2011;21(2):295–315.

10. Aubry T, Bell M, Ecker J, Goering P. Screening for housing first. *Canadian Observatory on Homelessness, Mental Health Commission of Canada*. 2015. https://www.homelesshub.ca/resource/screening-housing-first-phase-one-assessment-road-map.

11. Vaithianathan, Rhema, and Chamari I. Kithulgoda. Using predictive risk modeling to prioritize services for people experiencing homelessness in Allegheny County: Methodology update. 2020.

12. Azizi MJ, Vayanos P, Wilder B, Rice E, Tambe M. Designing fair, efficient, and interpretable policies for prioritizing homeless youth for housing resources. *Lect Notes Comput Sc*. 2018;10848 LNCS:35-51. doi:10.1007/978-3-319-93031-2_3

13. Amulya Y, Bryan W, Eric R, et al. Bridging the gap between theory and practice in influence maximization: Raising awareness about HIV among homeless youth. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*. 2018; 5399-5403. doi:10.24963/ijcai.2018/761

14. Chouldechova A, Putnam-Hornstein E, Benavides-Prado D, Fialko O, Vaithianathan R. A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. *Proceedings of Machine Learning Research*. 2018;81:1-15.

15. Gillman MW, Hammond RA. Precision treatment and precision prevention: Integrating "Below and Above the Skin." *Jama Pediatr*. 2015;170(1):1-2. doi:10.1001/jamapediatrics.2015.2786

16. Dolley S. Big data's role in precision public health. *Frontiers in public health.* 2018. https://doi.org/10.3389/fpubh.2018.00068

17. Dowell SF, Blazes D, Desmond-Hellmann S. Four steps to precision public health. *Nature*. 2016;540(7632):189–191.

18. Khoury MJ, Iademarco MF, Riley WT. Precision public health for the era of precision medicine. *Am J Prev Med.* 2016;50(3):398–401.

19. Hill JL. Bayesian nonparametric modeling for causal inference. *J Comput Graph Stat*. 2011;20(1):217-240. doi:10.1198/jcgs.2010.08162

20. Rahmattalabi A, Vayanos P, Dullerud K, Rice E. Learning resource allocation policies from observational data with an application to homeless services delivery. *arXiv preprint arXiv:2201 10053*. 2022.

21. Das, S. Local justice and the algorithmic allocation of scarce societal resources. *Proceedings of the AAAI Conference on Artificial Intelligence, 2022;36(11)*, 12250-12255. https://doi.org/10.1609/aaai.v36i11.21486

22. Tighe JR, Ganning JP. The divergent city: unequal and uneven development in St. Louis. *Urban Geogr*. 2015;36(5):654-673. doi:10.1080/02723638.2015.1014673

23. Arroyo-Johnson C, Woodward K, Milam L, Ackermann N, Komaie G, Goodman MS, Hipp JA. Still separate, still unequal: Social determinants of playground safety and proximity disparities in St. Louis. *J Urban Health*. 2016;93(4):627–638.

24. Murphy M. For the greater good of whom? Race, development, and civic values in the midwestern metropolis. *J Plan Hist*. 2006;5(4):355–364.

25. Chipman HA, George EI, McCulloch RE. BART: Bayesian additive regression trees. *Ann Appl Stat*. 2010;4(1):266-298. doi:10.1214/09-aoas285

26. Khadduri J, Leopold J, Sokol B, Spellman B. Costs associated with first-time homelessness for families and individuals. *HUD*. 2010. https://www.huduser.gov/portal/publications/povsoc/cost_homelessness.html

27. Culhane DP, Park JM, Metraux S. The patterns and costs of services use among homeless families. *Journal of Community Psychol*. 2011;39(7):815–825.

28. Gubits D, Shinn M, Bell S. Family options study: Short-term impacts of housing and services interventions for homeless families. *HUD, Office of Policy Development and Research*. 2015.

29. Elster J. *Local Justice: How Institutions Allocate Scarce Goods and Necessary Burdens*. Russell Sage Foundation; 1992.

30. Kyung L Min, Daniel K, Anson K, et al. WeBuildAI: Participatory framework for fair and efficient algorithmic governance. *Proceedings of the ACM Conference on Human-Computer Interaction.* 2019.

31. Karusala N, Wilson J, Vayanos P, Rice E. Street-level realities of data practices in homeless services provision. *Proceedings of the ACM Conference on Human-Computer Interaction*. 2019.

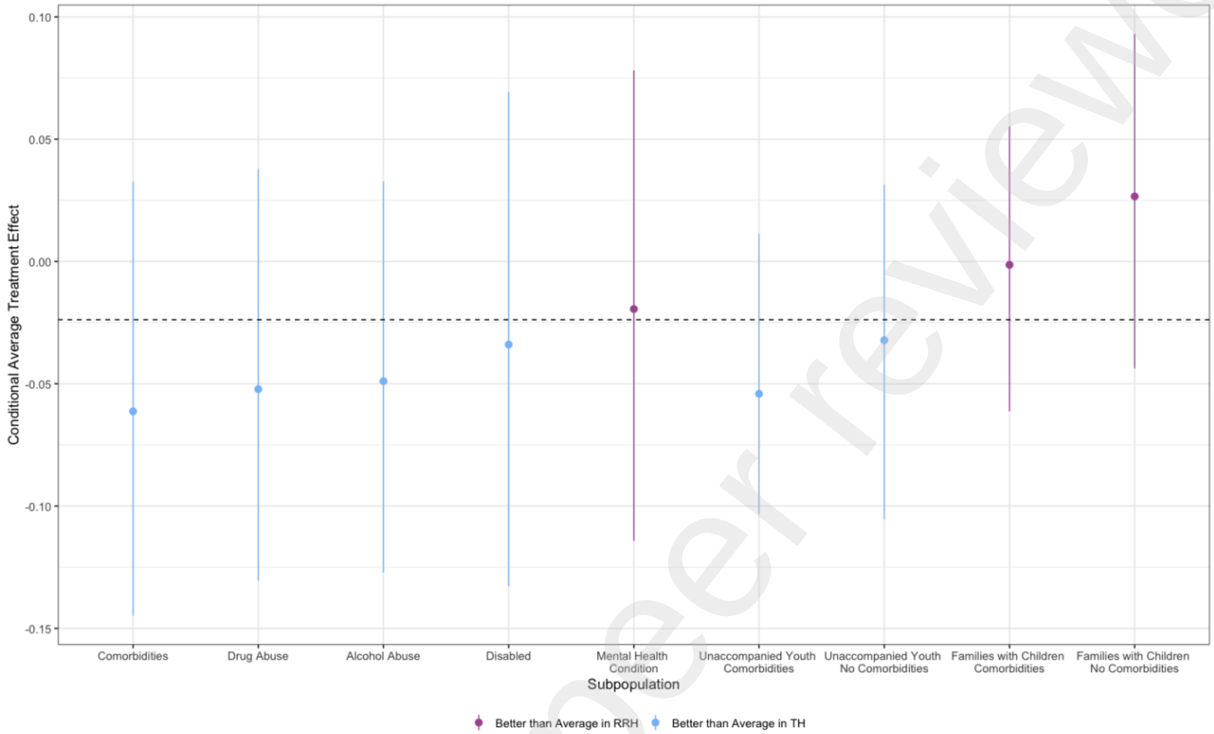
32. Vayanos P, McElfresh D, Ye Y, Dickerson J, Rice E. Active preference elicitation via adjustable robust optimization. *arXiv preprint arXiv:2003 01899*. 2020.

33. Kube A, Das S, Fowler P, Vorobeychik Y. Just resource allocation? How algorithmic predictions and human notions of justice interact. *Proceedings of the ACM Conference on Economics and Computation.* 2022.

Table 1. Allocation Comparison: Expected cost and reentry percentages for the decision rule and budget allocations compared to the original and unconstrained allocations

Allocation	Estimated Cost	Estimated Savings	(Expected) Reentry Percentage
Services-As-Usual	$5,468,446·10	-	27·82%
Prioritization Rules	$5,468,137·70	$308·40	26·78%
Cost-Effectiveness	$5,468,359·42	$86·68	26·17%
Service Efficiency	$6,838,410	$-1,369,963·90	25·61%

Figure 1. Conditional Average Treatment Effects of Transitional Housing Versus Rapid Rehousing by Subpopulation[a]



[a]The dotted line represents the ATE of transitional housing vs rapid rehousing for the entire population. CATE of zero indicate no treatment effect. CATE above the dotted line indicate the population performing better than average in rapid rehousing compared to TH; CATE below the dotted line indicate the population performing better than average in transitional housing. Bayesian credible intervals (and not confidence intervals) show the probability that each estimate falls within the 95% range

Figure 2. Prioritization Rules For Community- and Data-Driven Homeless Services

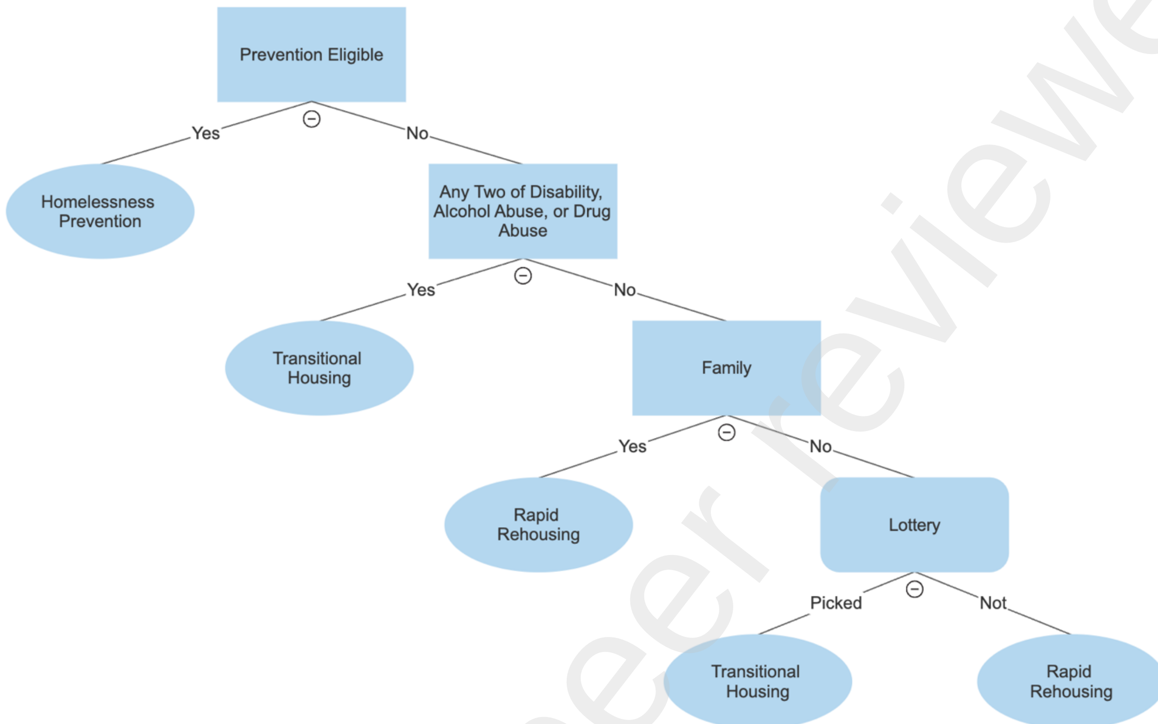
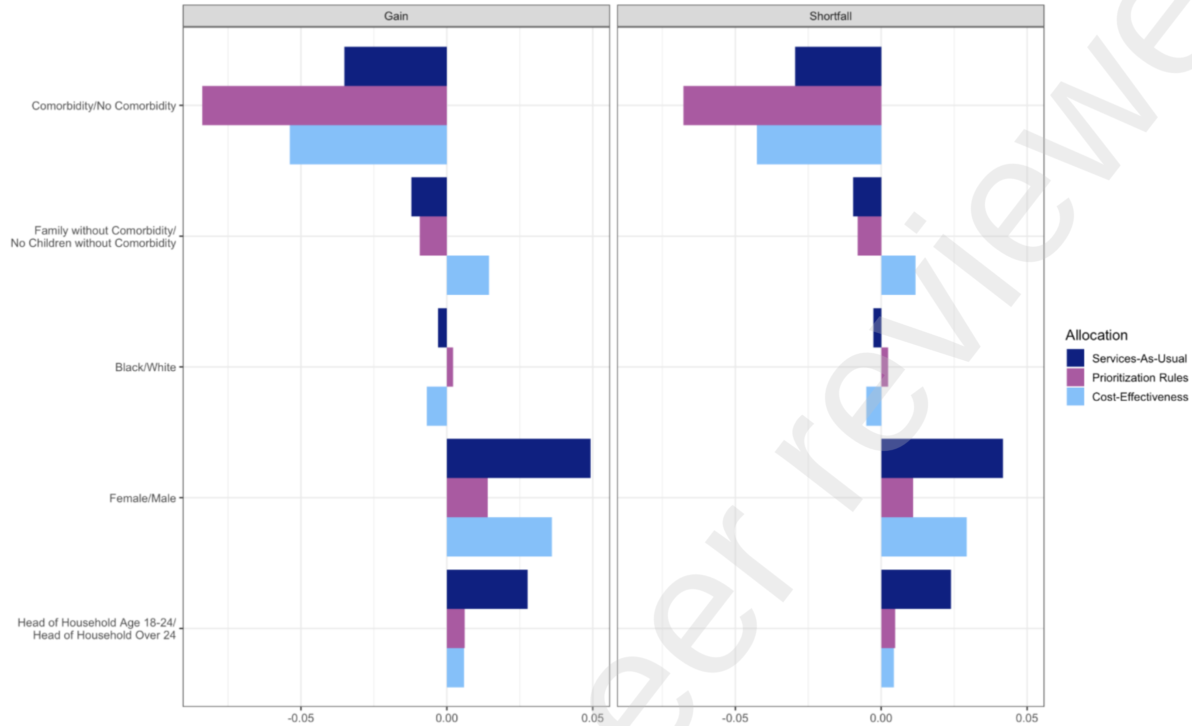


Figure 3. Group Fairness Bar Chart[a] Comparing Allocations on Equity Gains[b] and Shortfalls[c]



[a]Bars pointing to the left of the figure indicate bias toward Group A and bars pointing to the right of the figure indicate bias toward Group B where groups are listed as Group A/Group B

[b]Gain compares the utility with a worst-case scenario of serving all households with the service predicted to prevent homelessness the least, and thus, measures the benefit gained from each allocation.

[c]Shortfall contrasts with the best-case scenario when all households receive the service most likely to prevent homelessness – assessing the loss generated by each allocation.