Fair and Efficient Allocation of Scarce Resources Based on Predicted Outcomes: Implications for Homeless Service Delivery

Amanda R. Kube amanda.kube@wustl.edu
Division of Computational and Data Sciences, Washington University in St. Louis
1 Brookings Drive, St. Louis, MO 63130 USA

Sanmay Das sanmay@gmu.edu
Department of Computer Science
George Mason University
4400 University Drive, Fairfax VA 22030 USA

Patrick J. Fowler pjfowler@wustl.edu
Brown School of Social Work, Public Health, and Social Policy, and Division of Computational and Data Sciences
Washington University in St. Louis
1 Brookings Drive, St. Louis, MO 63130 USA

Abstract

Artificial intelligence, machine learning, and algorithmic techniques in general, provide two crucial abilities with the potential to improve decision-making in the context of allocation of scarce societal resources. They have the ability to flexibly and accurately model treatment response at the individual level, potentially allowing us to better match available resources to individuals. In addition, they have the ability to reason simultaneously about the effects of matching sets of scarce resources to populations of individuals. In this work, we leverage these abilities to study algorithmic allocation of scarce societal resources in the context of homelessness. In communities throughout the United States, there is constant demand for an array of homeless services intended to address different levels of need. Allocations of housing services must match households to appropriate services that continuously fluctuate in availability, while inefficiencies in allocation could "waste" scarce resources as households will remain needy and reenter the homeless system, increasing the overall demand for homeless services. This complex allocation problem introduces novel technical and ethical challenges. Using administrative data from a regional homeless system, we formulate the problem of "optimal" allocation of resources given data on households with need for homeless services. The optimization problem aims to allocate available resources such that predicted probabilities of household reentry are minimized. The key element of this work is its use of a counterfactual prediction approach that predicts household probabilities of reentry into homeless services if assigned to each service. Through these counterfactual predictions, we find that this approach has the potential to improve the efficiency of the homeless system by reducing re-entry, and, therefore, system-wide demand. However, efficiency comes with trade-offs - a significant fraction of households are assigned to services that increase probability of re-entry. To address this issue as well as the inherent fairness considerations present in any context where there are insufficient resources to meet demand, we discuss the efficiency, equity, and fairness issues that arise in our work and consider potential implications for homeless policies.

1. Introduction

Homelessness represents a long-standing social problem with considerable individual and collective costs. Homeless services coordinated at the community level (i.e, local homeless systems) have limited resources and therefore struggle to keep up with demand for housing assistance, and there is little evidence to support the efficiency of current decision making in the allocation of limited housing services as the allocation decisions themselves are under-studied (Brown, Cummings, Lyons, Carrión, & Watson, 2018; Fowler, Wright, Marcal, Ballard, & Hovmand, 2019b; Shinn, Greer, Bainbridge, Kwon, & Zuiderveen, 2013). Advances in machine learning and AI techniques have made it possible to apply learning algorithms to generate possible solutions to social problems ranging from raising HIV awareness (Yadav, Chan, Xin Jiang, Xu, Rice, & Tambe, 2016) to wildlife conservation (Dilkina & Gomes, 2010). In this paper, we explore the feasibility of data-driven approaches to inform policies that guide homeless service delivery. Specifically, we ask the question of whether individual predictions of success for certain types of homeless services can be leveraged to reduce the rate of re-entry into the homeless system across the population of households seeking assistance.

Background on resource allocation for social services: Public systems that coordinate responses to social problems face unique challenges at the intersection of efficiency and fairness. Social services aim to address a wide array of homeless household needs through a host of services that range in time and intensity; providers continuously make decisions on whom to serve with what service or combination of services. Moreover, constant resource constraints limit the capacity to address widespread demand for assistance. The information available for decision-making is far from perfect given the imprecision of needs assessments, as well as poor understanding of what services work for whom (Gubits, Shinn, Wood, Brown, Dastrup, & Bell, 2018; Shinn, Brown, Spellman, Wood, Gubits, & Khadduri, 2017). In the context of scarcity, providers make complex decisions under great uncertainty with small margins of error. Poor decisions that either under- or over-serve homeless households waste scarce resources and miss opportunities for meeting the needs of those not served at all.

In the algorithmic decision-making literature on social service provision, the typical approach is to prioritize decisions based on risk scores. For example, Chouldechova and colleagues consider risk assessment in the context of child maltreatment to decide on which calls to a child protection hotline should be investigated further (Chouldechova, Benavides-Prado, Fialko, & Vaithianathan, 2018; Brown, Chouldechova, Putnam-Hornstein, Tobin, & Vaithianathan, 2019). These represent classic triage situations, and deal with the problem of which cases to select given a limited budget and a risk assessment. Another context of algorithmic decision-making concerns online resource allocation – when a resource becomes available, which of various agents waiting in a queue should be allocated that resource? The most relevant study along these lines is that of Azizi, Vayanos, Wilder, Rice, and Tambe (2018), who consider allocation policies specifically for homeless youth. They formulate a dynamic allocation problem between arriving homeless youth and two types of housing resources (rapid rehousing and permanent supportive housing) and consider the issues involved in fair and efficient online allocation of youth to these resources.

The market and mechanism design literature have conducted considerable research on assignment problems, school allocation, organ allocation, refugee matching, etc. (Kominers, Teytelboym, and Crawford (2017) provide an excellent recent introduction to market design). A key focus there has been on the preferences and incentives of the participants, as well as the level of control of the mechanism in allocation decisions. For example, in traditional assignment problems, it is assumed that the principal has full control over all allocation decisions (Kuhn, 1955); much of the literature on two-sided matching seeks stable matchings that respect pairwise preferences (Roth & Sotomayor, 1990); work on school choice considers student preferences and school priorities differently (Abdulkadiroglu, Pathak, Roth, & Sonmez, 2005); the kidney exchange literature seeks to maximize the number of matches of incompatible pairs (Roth, Sonmez, & Utku Unver, 2005; Dickerson, Manlove, Plaut, Sandholm, & Trimble, 2016).

One of the main benefits of our approach is the possibility of increasing efficiency by exploiting gains from heterogeneity in match quality between households and services. This issue has been explored infrequently in the market design literature, perhaps because of the historical focus on ordinal preferences rather than cardinal utilities (Anshelevich & Das, 2010; Anshelevich, Das, & Naamad, 2013), which better aligns with systems where agents have considerable control in terms of accepting and rejecting their assignment or matching. However, consideration of cardinal utilities has come up recently in the context of compatible living donor kidney transplantation (Li, Lieberman, Macke, Carrillo, Ho, Wellen, & Das, 2019) where one can take advantage of differences in match quality between the organ and the patient and in refugee matching (Bansak, Ferwerda, Hainmueller, Dillon, Hangartner, Lawrence, & Weinstein, 2018; Trapp, Teytelboym, Martinello, Andersson, & Ahani, 2021) where one can optimize over utilities of matchings between refugees and resettlement venues.

Recent approaches to refugee matching, roughly contemporaneous with this research, are the closest point of comparison to our work. Trapp et al. (2021) use a combination of machine learning and integer programming to optimize employment outcomes for resettled refugees using historical data. Following Bansak et al. (2018), Trapp et al. (2021) take advantage of the randomness present in current refugee assignment to ensure that selection bias does not affect their modeling. Though we also use machine learning and integer programming on historical data, our setting offers a different challenge given that housing services in our data are not assigned to households at random. Therefore, our observational data, which is routinely collected as part of service provision, is confounded by caseworker decisions. This magnifies the importance of causal modeling. As opposed to the types of problems that Kleinberg, Ludwig, Mullainathan, and Obermeyer (2015) call "prediction policy problems", or for example using machine learning predictions of loan default to manage risk (Butaru, Chen, Clark, Das, Lo, & Siddique, 2016), we need useful counterfactual estimates of the effects of different services in order to begin defining the resource allocation problem.

There has been significant recent progress in causal modeling from a machine learning perspective (Johansson, Shalit, & Sontag, 2016, e.g.). Bayesian counterfactual approaches offer particular promise for informing social services delivery (Hill, 2011), since Bayesian models can provide coherent probabilistic estimates of heterogeneous treatment effects, and thus, allow predictions of individual outcomes under counterfactual predictions (Chipman, George, & McCulloch, 2007; Chipman, George, McCulloch, et al., 2010). Therefore, we test a promising Bayesian model for this purpose. While our primary goal is to use it in the counterfactual prediction setting, we also compare it with other learning algorithms using standard machine learning metrics on the typical out-of-sample prediction task.

Ours is one of the first studies to consider using machine learning-based estimates of counterfactual outcome probabilities to estimate the value of, and thus inform, allocation decisions for social services, specifically interventions for homeless households. We present this work as a proof-of-concept, based on a real administrative dataset across the whole range of homeless populations in a metro area, to address the following question: *By optimizing allocations based on predicted outcomes, how much could we potentially improve outcomes, and what would be the distributional effects of these improvements?*

Problem setup: Local homeless systems coordinate community-wide services that address housing crises. In the US, services range in intensity from time-limited nonresidential supports to ongoing rental assistance with intensive case management (United States Congress, 2009). Each service is capacity constrained, given the constant widespread demand for affordable housing. Thus, homeless providers allocate many households to many services that each vary in availability at any given time. Homeless services aim to stabilize households and reduce future demand for assistance.

National policies currently focus evaluation of homeless service delivery on whether households use additional homeless services within two years of entry into the system; counts are generated from administrative data that record entries and exists across homeless services (HUD, 2012). However, routine capacity constraints make it challenging to measure success, since those in need may not be able to receive services. Missing information impedes service improvements in most communities across the US (Fowler et al., 2019b).

In this work, we take advantage of unique local administrative records that capture community-wide demand and receipt of homeless assistance across time. The data we use link homeless service records with *requests* for assistance through a regional homeless hotline. Operators at the central hotline field all requests for services, as well as make referrals to appropriate and available services. Households call back if they are in need of additional services, and a digital trail captures subsequent requests, regardless of eligibility or delivery of services. This extensive data collection exceeds federal requirements and allows for a comprehensive assessment of homeless services impossible for most communities.

We test several machine learning methods for predicting outcomes of matching households to different interventions, and choose to use BART (Bayesian Additive Regression Trees) because it is competitive in out-of-sample performance with other methods, allows for meaningful probability distributions over outcomes, and has been established as a powerful method for causal inference with observational and complex data (Hill, 2011). Using BART, we build and evaluate counterfactual models for whether a household would have re-entered the homeless system within 2 years if they had been assigned to a different service, and solve a capacitated assignment problem in order to minimize the number of households re-entering the system within two years, subject to capacity constraints on each service.

Preview of results: Using administrative data on a weekly basis over the course of 166 weeks, we estimate counterfactual predictions of reentry into homeless services for each household within two years. Models appear well-calibrated; we predict (out-of-sample), in expectation, that 2725 (27.13%) households would re-enter the system, whereas 2765 (27.53%) actually re-entered. In the optimized assignment, we find the BART model predicts that only 2155 households (21.46%) would re-enter the system. Thus, there may be substantial benefits achievable (by this re-entry metric) from improving the combined prediction-allocation mechanism. However, these benefits come with tradeoffs. The allocations are not Pareto-improving; nearly one-third of households increase their probability of re-entry according to the predictions. We formulate and solve a constrained version of the allocation problem that guarantees no household increases the probability of re-entry by more than 5 percentage points in the new allocation. In this case, 22.29% of households are predicted to re-enter.

Implications: Our work serves as a proof of concept through a case study. We bring administrative data to bear on the question of how much AI techniques can improve social service provision, with full awareness that the precise results presented may depend on specific modeling choices, and the reliability of the counterfactual estimates. This work contributes to the emerging dialogue on social service delivery based on machine learning predictions. We emphasize the importance of considering fairness, ethics, and the long-term dynamics of systems that use these kinds of predictive models, while at the same time believing that engaging these questions with actual data and estimates can contribute to resolving the lack of evidence guiding current social service delivery.

2. Ethics and fairness:

Since we are considering a problem of allocating scarce, shared societal resources using algorithmic approaches, it is important to foreground the discussion of ethical issues and fairness concerns. The use of techniques from machine learning and artificial intelligence (and more broadly, algorithmic approaches) in different societal contexts increasingly raise concerns regarding fairness, accountability, and transparency (O'Neil, 2016, among others). Although they demonstrate potential for improvements in efficiency, fundamental questions exist as to whether data-driven allocations introduce or perpetuate systematic biases that contribute to inequities, and moreover, whether the inherent complexity of decision-making impedes timely detection and correction of these inequities. A number of recent studies justify these concerns, demonstrating racial disparities in credit lending, hotspot policing, and crime sentencing (Ensign, Friedler, Neville, Scheidegger, & Venkatasubramanian, 2018; Pleiss, Raghavan, Wu, Kleinberg, & Weinberger, 2017; Corbett-Davies, Pierson, Feller, Goel, & Huq, 2017); each example shows that marginalized, underrepresented minorities disproportionately suffer from unfair algorithmic decisions. The unintended consequences require that we carefully consider how to design adequate protections against systematic misuses.

The European Union recently passed legislation in response to concerns about ethics, fairness, and privacy. The "General Data Protection Regulation" (GDPR) imposes restrictions on how individual data can be used for algorithmic decision making in ways that "significantly affect" users. The GDPR coincides with a broader argument for not just full transparency, but rather *human interpretability* regarding how decisions are derived from algorithmic approaches to ensure adequate assessment of fairness. However, requirements for human interpretability could also diminish the potential of AI to solve societal problems. Algorithmic approaches generate novel solutions that may not correspond to human intuition; requirements for full explainability of these complex processes limits the inherent value of applications to thorny social problems.

David Weinberger presents a compelling example related to autonomous vehicles in a *Wired* op-ed (Weinberger, 2018). If self-driving automobiles lowered the number of vehicular fatalities by 90%, would it really be worth losing that benefit because of the difficulty of explaining (or legal liabilities that may be associated with) the remaining crashes? Certainly, the answer partly depends on whether the remaining crashes disproportionately affect some portion of the population as well as other considerations. Weinberger goes on to argue that while the regulation of AI applied to social problems is critical, it can be achieved through existing processes for resolving policy issues (Weinberger, 2018). Governance provides formal and informal methods for establishing rules and norms applied to collective problems, which also include sustainable approaches for mutual accountability. According to Weinberger, the right approach towards AI regulation involves specification of appropriate optimization goals arrived through the social processes of policy-making that consider both efficiency and equity. However, with a few exceptions (Chouldechova et al., 2018, e.g.) there has not been much empirical investigation probing the tradeoffs that emerge when incorporating fairness considerations into algorithmic decisions, especially in the context of scarcity.

3. Background and Data

Homelessness represents a complex public health challenge for communities across the United States. Federal guidelines define homelessness as residence in unstable and non-permanent accommodations. This includes shelters, places not meant for habitation (eg., cars, park, abandoned buildings), as well as being at imminent risk for eviction. Annual counts since 2007 estimate that more than 550,000 people experience homelessness on a single January night across the United States (Henry, Mahathey, Morrill, Robinson, Shivji, & Watt, 2018), while approximately 1.5 million people use homeless services at some point during each year (Henry et al., 2018). Families with children under 18 years of age comprise more than one-third of homeless households (Henry et al., 2018). Experiences of homelessness and associated turmoil carries life long implications, as well as significant social costs in lost productivity, compromised health, and compensatory social service expenditures (Khadduri, Leopold, Sokol, & Spellman, 2010; Culhane, Park, & Metraux, 2011; Fowler, Hovmand, Marcal, & Das, 2019a).

The homeless system represents the primary community-wide service response to housing crises. Funds allocated by Congress on an annual basis support the delivery of five types of homeless assistance. Service types vary in intensity, and relatedly, availability. The most intensive service - Permanent Supportive Housing - provides long-term rental assistance plus comprehensive case management to address barriers to stability, such as mental health and substance abuse treatment; it is reserved for the highest risk households and consumes the greatest amount of financial resources. Similarly to Permanent Supportive Housing, Transitional Housing also offers comprehensive case management but only up to 24 months in congregate settings. Rapid Rehousing allows up to 24 months of rental assistance without additional intensive case management. At the end of two years, households in Transitional Housing or Rapid Rehousing either move on their own or step-up to Permanent Supportive

Housing, if available. Emergency Shelters offer immediate accommodations for those with no other place to go, and typically serve a large number of households for a brief period of time. Shelters are intended to stabilize households and divert high-risk families to the longer-term housing services. Finally, Homelessness Prevention provides households at imminent risk for homelessness with short-term and non-reoccurring assistance to mitigate housing crises. Local non-profit provider networks determine the delivery of day-to-day services within general structures determined by federal funding priorities. During the study period, providers offered services to eligible households on a first-come-first-served basis.

Despite substantial investments, homeless rates remain high in the United States (Fowler et al., 2019a). An enormous challenge is that of matching service types to need. While federal guidelines mandate that local agencies provide services based on risk assessments (United States Congress, 2009), existing tools fail to discern high and low risk households reliably and accurately (Brown et al., 2018; Shinn et al., 2013). Homeless service providers have limited evidence for adapting responses to observed and unobserved household characteristics (Fowler et al., 2019b). Moreover, there are no tools that assess the impact of service matches on overall system performance in reducing reentries.[1]

3.1 Data Collection

Data for this work come from the homeless management information system (HMIS) of a major metropolitan area from 2007 through 2014. The HMIS records all housing services provided to individuals and families seeking federally funded homelessness assistance. Local service providers enter information on requests and receipt of services in real time through a web-based platform in accordance with federal mandates for collection of universal elements. A local non-profit organization contracted with the homeless system hosts the platform and provides support, including user training, technical assistance, and active quality control.

Records provide information on the characteristics and services delivered to households in contact with the homeless system. Household-level characteristics include an array of information on demographics, housing risk, and eligibility determinations. Services include entry and exit dates from the five federally defined types of homeless assistance: homelessness prevention, emergency shelter, rapid rehousing, transitional housing, and permanent supportive housing. In addition, the metropolitan area coordinates requests for assistance through a homeless hotline, and household-level data record information on every call, including dates and referral for services. Household identifiers allow linkages of information across time. Data sharing agreements with regional homeless systems allow access to deidentified records in accordance with the relevant Institutional Review Board, which made a non-human subjects determination. Regardless, all information was transferred, stored, and analyzed according to best practices in data security. This includes ethics training in research for all research team members.

1. Annual evaluations of homeless system performance monitor overall rates of return to the homeless system within 24 months, but do not evaluate allocations; future federal funding depends in part on demonstrating trends toward reductions in reentries.

3.2 Data Cleaning and Feature Selection

For this project, we extract data provided by 75 different homeless agencies and link participants across programs by a unique, anonymous identification number. We then aggregate data by household over time using a unique household identification number. This results in a dataset of households containing household characteristics available upon entry into the system, as well as information on all entries and exits from different homeless services. We exclude permanent supportive housing for the present study because the service was rarely used as an initial response for first time entries into the homeless system during the study period. The primary outcome (the label we are trying to predict) is reentry into the homeless system. Operationally, reentry is defined as requesting services within two years of exit from the system, regardless of whether services were actually received. We do this using hotline call records to determine whether a household requested additional housing assistance after the initial service. This ensures that we capture further need, and not just availability of services. When transitions between services (e.g. homeless shelter to rapid rehousing) occur on the same day, we assume that they represent a continuation of homeless services and do not count this as a reentry. We consider households to have exited from the system when the time between leaving one service and entering another exceeds one day. Our analyses include households who entered the homeless system after the start of 2007 and exited before the end of 2012 to provide a minimum two-year follow-up for all households.

Type	Number	Examples
Binary Features	3	Gender, Spouse Present, HUD Chronic Homeless
Non-Binary Categorical Features	19	Veteran Status, Disabling Condition, Substance Abuse
Continuous Features	13	Age, Monthly Income, Calls to Hotline, Duration of Wait
Total Features	35	

Table 1: Summary of features included in BART model

Since the data captures homeless services across time, it contains both time-invariant (e.g., race, gender, ethnicity) as well as time-variant (e.g., monthly income, age) features. We select values of time-variant features that are collected at the time of first entry into the homeless system and have adequate amounts of available data for use in modeling. Most of the variables we selected were categorical, and missing values are treated as a separate category in these cases. Table 1 shows a summary and examples of the features included. A more complete summary of the dataset is included in Table 8 in the Appendix.

3.3 Data Characteristics

The dataset includes records on 13940 households. The target variable, or label, is a binary indicator of whether households reentered the homeless system, defined as requesting and/or receiving homeless services within 2 years of initial exit. Of the 13940 households, 3987 (28.60%) reentered the homeless system within two years; among reentries, 2066 (51.82%) were placed in a subsequent service, while 1921 (48.18%) called the hotline to request services, but by the end of the two year period had not been placed in another service. Reasons for failing to receive additional services varied; most commonly, services were unavailable and clients were referred to other services (79.13%) or clients did not follow up on referrals (17.67%).

Table 2 shows the number of households initially assigned to each homeless service type, as well as the percentage of reentries within 2 years for each service. Models use a single feature vector, which consists of service assignment plus additional covariate data collected at first entry into the system.

Service Type	Number Assigned	Percent Reentered
Emergency Shelter	4431	43.11
Transitional Housing	2449	34.38
Rapid Rehousing	844	40.40
Homelessness Prevention	6216	14.38
Total	13940	28.60

Table 2: Summary of service assignment and homeless system reentry within two years by type of service

4. Analyzing Services

This application requires a method that can handle the challenges of counterfactual inference using observational data, while simultaneously providing a well-grounded probabilistic model. BART, an ensemble model that outperforms propensity score and nearest neighbor matching algorithms for causal inference on observational data, especially when the data are complex (Hill, 2011), is a promising method for mitigating this challenge, (Chipman et al., 2007, 2010).

Bayesian nonparametric modeling for causal inference has a number of advantages that fit this application (Chipman et al., 2010; Hill, 2011; Johansson et al., 2016). Such models are capable of providing robust estimates of treatment effects using observational data like administrative service records. They can handle a large number of features or predictors, as well as complex data that include interactions and nonlinearities seen in prior studies of homeless service delivery (Shinn et al., 2013). In the following section, we compare the predictive performance of BART on our dataset to that of several other popular machine learning algorithms: random forests, logistic regression, LASSO, and gradient boosted trees.

4.1 Model Comparison

We compared the out-of-sample predictive performance of BART to four commonly used machine learning algorithms using 10-fold cross validation. First, we implemented BART using the default parameters provided by the model creators (Chipman et al., 2010). Then, we implemented simple logistic regression and LASSO using 10-fold cross validation to choose the value of lambda, the regularization parameter. We also implemented random forests with 500 trees, a minimum node size of 1, and considering 6 variables for each split (Breiman, 2001). Lastly, we implemented gradient boosted trees with 100 trees with a maximum depth of 1, 10 observations per node minimum, and a learning rate of 0.1. As we are comparing to BART using the default parameters, these hyperparameters were chosen because they are commonly used default parameters/implementations for each method. We assess predictive performance using multiple metrics: AUC (Area Under the ROC Curve), Misclassification Error, Precision, Recall, and Calibration which we operationalize as Expected Reentries/True Reentries. We also assess calibration individually for each service type operationalized in the same manner. The results of this analysis are shown in Tables 3 and 4. BART outperforms each of the other methods and, as stated previously, mitigates the issue of confounder bias that may be present in our observational data and allows for the estimation of household-specific treatment effects. For these reasons we chose to conduct all future analyses using BART. All model fitting and counterfactual inference that follows is done using the R package `BayesTree` written by the model's creators (Chipman et al., 2010).

Method	AUC	Misclassification Error	Precision	Recall	Calibration
BART	0.7534	0.2506	0.6136	0.3393	0.9999
Logistic Regression	0.7386	0.2576	0.6171	0.2670	0.9996
LASSO	0.7386	0.2583	0.6254	0.2465	0.9995
Random Forests	0.7444	0.2516	0.6110	0.3361	0.8864
Gradient Boosted Trees	0.7462	0.2564	0.6104	0.2920	0.9999

Table 3: Comparison of prediction performance of several commonly used methods using multiple metrics

Method	Emergency Shelter	Transitional Housing	Rapid Rehousing	Homelessness Prevention
BART	0.9990	1.0009	0.9961	1.0022
Logistic Regression	1.0001	0.9989	0.9980	0.9999
LASSO	0.9921	1.0059	0.9753	1.0183
Random Forests	0.9423	0.8285	0.9824	0.7860
Gradient Boosted Trees	0.9747	1.0371	0.9382	1.0419

Table 4: Comparison of the calibration of each method by service type

4.2 Counterfactual Estimation of Heterogeneity in Match Quality

Using BART, we built models to produce out-of-sample counterfactual estimates of reentry probabilities if households received each homeless service (i.e., prevention, rapid rehousing, shelter, transitional housing).[2] For most of the 13940 households, either homelessness prevention or transitional housing produce the lowest probability of reentering the system within two years (9152 households are predicted to do best in prevention and 4907 households in transitional housing). Four households were predicted to do best in emergency shelter and 75 in rapid rehousing. Most households were predicted to have the highest probability of reentry if placed in emergency shelter (6223 households) or transitional housing (6003 households) with less predicted to do worst in rapid rehousing and prevention (1688 households and 26 households respectively).

2. These counterfactual estimates for all 13940 households are made available in the following repository: https://github.com/amandakube/Allocating-Homelessness-Interventions—Counterfactual-Predictions

Relative Ordering of Services	Number of Households	Average Probability of Reentry in ES	Average Probability of Reentry in TH	Average Probability of Reentry in RRH	Average Probability of Reentry in Prev
Prevention, Rapid Rehousing, Emergency Shelter, Transitional Housing	4367	0.25	0.41	0.23	0.16
Transitional Housing, Prevention,Rapid Rehousing, Emergency Shelter	3437	0.47	0.21	0.43	0.36
Prevention, Transitional Housing, Rapid Rehousing, Emergency Shelter	1806	0.30	0.24	0.27	0.19
Prevention, Emergency Shelter, Rapid Rehousing, Transitional Housing	1506	0.21	0.38	0.22	0.14
Transitional Housing, Prevention, Emergency Shelter, Rapid Rehousing	1006	0.47	0.21	0.50	0.39
Prevention, Rapid Rehousing, Transitional Housing, Emergency Shelter	751	0.29	0.27	0.25	0.19
Prevention, Transitional Housing, Emergency Shelter, Rapid Rehousing	548	0.27	0.23	0.29	0.19
Transitional Housing, Rapid Rehousing, Prevention, Emergency Shelter	215	0.58	0.27	0.51	0.53
Prevention, Emergency Shelter, Transitional Housing, Rapid Rehousing	174	0.23	0.24	0.25	0.16
Rapid Rehousing, Prevention, Emergency Shelter, Transitional Housing	56	0.38	0.51	0.32	0.34
Transitional Housing, Emergency Shelter, Prevention Rapid Rehousing	25	0.60	0.30	0.65	0.61
Transitional Housing, Rapid Rehousing, Emergency Shelter, Prevention	15	0.47	0.22	0.43	0.47
Rapid Rehousing, Prevention, Transitional Housing, Emergency Shelter	12	0.41	0.39	0.34	0.36
Transitional Housing, Emergency Shelter, Rapid Rehousing, Prevention	11	0.51	0.19	0.52	0.53
Rapid Rehousing, Emergency Shelter, Prevention, Transitional Housing	5	0.34	0.44	0.33	0.35
Emergency Shelter, Prevention, Rapid Rehousing, Transitional Housing	3	0.32	0.64	0.37	0.33
Rapid Rehousing, Transitional Housing, Prevention, Emergency Shelter	2	0.43	0.40	0.36	0.40
Emergency Shelter, Prevention, Transitional Housing, Rapid Rehousing	1	0.13	0.15	0.15	0.14

Table 5: Number of households having each of the orderings of services from least to greatest probability of reentry

For each household, we determined which services are predicted to outperform others and developed a relative ordering of service effectiveness. Table 5 illustrates this ordering of service effectiveness. Summing across households, almost one-third (31.3%) do best in prevention followed by rapid rehousing, shelter, and transitional housing. Another 24.7% would benefit most in transitional housing, followed by prevention, rapid rehousing, and shelter. For a small proportion of households (13.9%), transitional housing followed by prevention, rapid rehousing, and shelter would be best. The remaining third of households would be best served by different orderings of these services. These patterns demonstrate the heterogeneity in treatment effects we hope to leverage to improve the efficiency of allocations.

The probabilities estimated by BART allow us to perform an initial examination of the possibility of optimizing homeless service delivery. If all households were placed in the service in which they have the lowest predicted probability of reentry, we predict 18.47% of households would reenter in expectation. This is a 35.42 percent decrease from the 28.60% who actually reentered. However, it represents a vast oversimplification of the allocation problem, which in reality is subject to capacity constraints on the number of households that can be served by a particular service at any given time. In the following section, we formulate the optimal allocation problem including these service capacity constraints.

5. Optimal Allocation Using Estimated Personalized Treatment Effects

In order to frame the optimal allocation problem, we need two main sets of variables estimated from the data. First are the actual predictions of probability of reentry for households given they are placed in each of the possible services. For this, we use out-of-sample BART predictions. Second are the capacities of the different services mentioned in the previous section - that is, the number of households that can be accommodated at a given time due to space or monetary limitations. In order to estimate these, we aggregate data on a weekly basis, and set the capacity of a service equal to the number of households who truly entered into the service in that week. One week is granular enough to give some flexibility to the optimizer, while also not leading to waits that are outside the tolerance of the system. We note here that we solve the problem in a static manner every week, although there could, of course, be interesting dynamic matching issues at play (Akbarpour, Li, & Gharan, 2020; Anshelevich, Chhabra, Das, & Gerrior, 2013).

5.1 The Optimization Problem

Let x_{ij} be a binary variable representing whether or not household i is placed in service j. Then, the Integer Programming problem is given by

$$\begin{aligned} \min_{x_{ij}} \sum_i \sum_j p_{ij} x_{ij} & \\ \text{subject to} \quad \sum_j x_{ij} = 1 & \quad \forall i \in \mathbb{Z} \\ \sum_i x_{ij} \leq C_j & \quad \forall j \in \mathbb{Z} \\ x_{ij} \in \{0, 1\} & \end{aligned}$$

where p_{ij} is the probability of household i reentering if they are placed in service j and C_j is the capacity of service j.

We use this IP framework and Gurobi optimization software to find an optimal allocation for households who entered the system during each week.

In the following section, we show this can be re-formulated as a weighted bipartite b-matching problem, known to admit a polynomial time solution.

5.1.1 Reduction to Weighted Bipartite b-matching

Weighted Bipartite b-Matching is the following problem: Given a weighted bipartite graph G with positive, real-valued edge weights, find a subgraph H of G with maximum total weight such that every vertex i in H is incident to at most b_i edges (Chen, Zheng, Srinivasan, Thomo, Wu, & Sukow, 2016).

Given an instance of the current optimization problem, we create an instance of Weighted Bipartite b-Matching as follows. First, create a bipartite graph G such that there are four nodes representing the four services on the right and a single node representing each household on the left. Between each household node i and each service node j, create an edge and give that edge weight $1 - p_{ij}$. For each household node i, let the degree constraint b_i of node i be 1. For each service node j, let the degree constraint b_j of node j be C_j. Then, the allocation of households to services that minimizes expected re-entries while respecting capacity constraints is given by a maximum weighted bipartite b-matching on graph G.

Claim: An optimal weighted bipartite b-matching solution of maximum weight on graph G gives an allocation of households to services that solves the current optimization problem.

Proof: Assume there exists an optimal weighted bipartite b-matching solution of maximum weight on graph G that does not give an allocation of households to services that minimizes expected re-entries while respecting capacity constraints.

We know that each household i is going to be matched to exactly one service j since each household node in G has capacity 1 and not fulfilling that capacity can only reduce the total weight of the solution. Similarly, each service j must be at capacity, since leaving any household unmatched would only result in a solution of smaller total weight. Therefore, if there is an improvement to be made to the optimal allocation of households to services, it must be due to swapping some pair of edges. Now, suppose household h is assigned to service k and h' to service k'. Suppose swapping them so that h were assigned to k' and h' to k would improve the re-entry minimization objective. Then it must be the case that $p_{hk} + p_{h'k'} > p_{hk'} + p_{h'k}$. Which implies $(1 - p_{hk}) + (1 - p_{h'k'}) < (1 - p_{hk'}) + (1 - p_{h'k})$. Therefore, swapping them would increase the total weight of the weighted bipartite b-matching solution.

This contradicts the assumption that our solution to the weighted bipartite b-matching problem was of maximum weight. Therefore, the allocation of households to services that minimizes expected re-entries while respecting capacity constraints must be given by a maximum weighted bipartite b-matching on graph G. ■

This shows that the solution to our optimization problem can be found in polynomial time. In practice, the optimization is extremely fast in Gurobi (0.03 seconds on average), and time requirements are dominated by running BART, therefore we use the MIP formulation.

5.1.2 Optimization Results

Only households who entered the homeless system between October, 2009 (after initial implementation of the rapid rehousing service) through December, 2012 were included in the optimization. This results in tracking 10043 households across 166 separate weeks optimized.[3]

Over the 166 weeks, 2765 out of 10043 households (27.53%) actually reentered the homeless system. Summing BART predictions to estimate how many households would reenter in expectation produces an estimate of 2725 households (27.13%), suggesting that the predicted reentry probabilities given by BART are reliable. Using these predicted probabilities to find an optimal allocation, predicted reentries reduce to 2155 households (21.46%). Thus, the optimal allocation framework reduces the predicted number of reentries into the homeless system by 22.06% over this period. Also recall that the best that could be achieved by assigning each household to its optimal service, without any capacity constraints, was a reentry rate of 17.97%, so our allocation gets us much closer to the best possible reentry rate for this formulation.

5.2 Fairness Considerations

An immediate question is whether the optimal allocation is capturing some inherent inefficiency in the allocation system, and is therefore Pareto-improving or at least improving allocations for a substantial portion of the population.

Figure 1 shows the distribution of changes in predicted probability of reentry based on our BART model in the optimal service versus predicted probability of reentry for the actual service allocation. In the optimal allocation, 3967 (39.50%) individual households are allocated to a service in which they have a lower probability of reentry than the service in which they actually participated (shown by the area of the histogram to the right of 0). Another 3251 (32.37%) are allocated to the same service they were originally assigned. Importantly, 2825 (28.13%) households are allocated to a service in which they have a *higher*

3. Two simultaneous changes in homeless service delivery precluded additional follow-up. First, new data management software failed to match households in the system before and after 2015. Second, local homeless providers simultaneously shifted services to comply with federal requirements for coordinated entry into homeless services; the result, in effect, unpaired prevention from other homeless services.

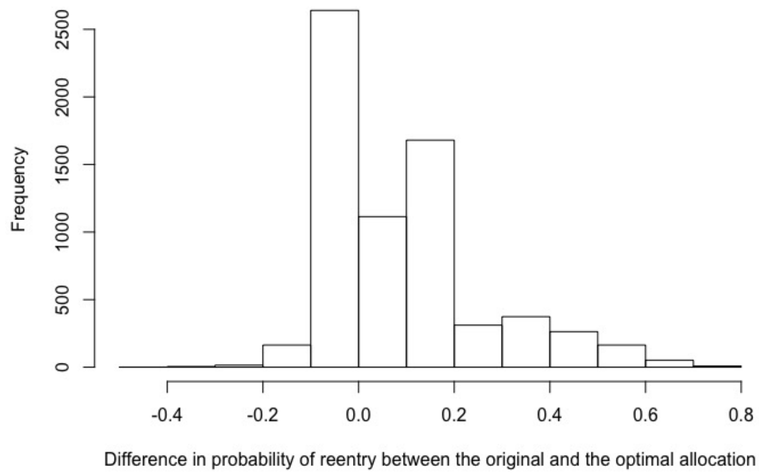


Figure 1: Histogram of improvement in reentry probability under the unconstrained optimized allocation (the 3251 households whose probability of reentry was unchanged are not included)

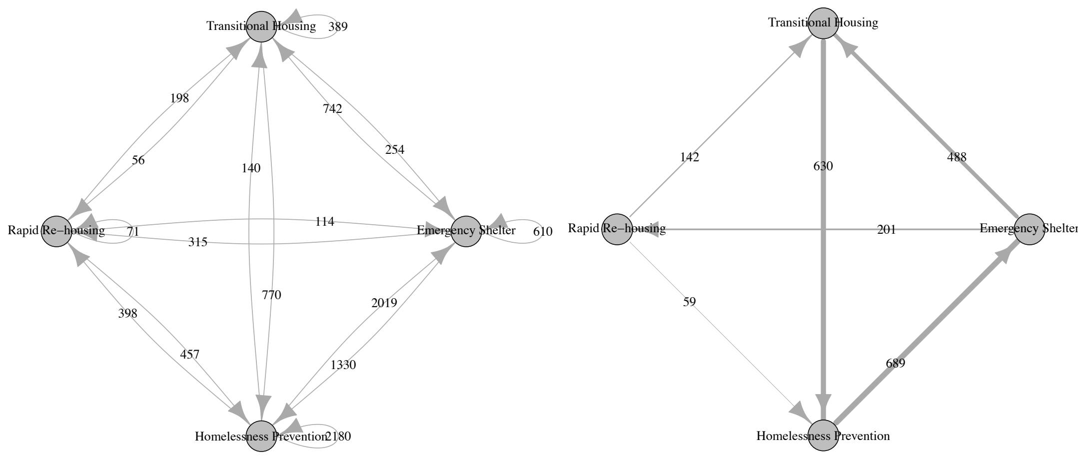
probability of reentry (shown by the area of the histogram to the left of 0). Therefore, a substantial fraction of households are being hurt by the reassignment, even though more are being helped.

Original \ Optimal	Emergency Shelter	Transitional Housing	Rapid Re-housing	Homelessness Prevention
Emergency Shelter	0	0.39	0.06	0.12
Transitional Housing	0.02	0	0.04	0.08
Rapid Re-housing	0.02	0.41	0	0.12
Homelessness Prevention	-0.06	0.21	-0.04	0

Table 6: Average percentage point difference in probability of reentry for households moving between services in the optimal and original allocations. Positive numbers represent decreases in probability of reentry.

Table 6 shows the average percentage point difference in probability of reentry for households moving from one service in the original allocation to a different service in the optimal allocation. The mainly positive non-zero off diagonals suggest potential improvements from optimization that range from small (e.g., rapid rehousing to shelter) to larger changes, especially reassignment to transitional housing. Although BART shows homelessness prevention represents the best option for most households, the percentage point gains are relatively modest; those who are moved out of prevention typically have worse outcomes.

Figure 2 shows the mechanism of improvement, given capacity constraints. It maps the changes in allocation between the different services in the optimal allocation, as compared with the original. Figure 2a shows the number of households who moved from each service to another in the optimal allocation and Figure 2b shows the net flows of households moving



(a) Full network map (b) Network map with net flow

Figure 2: Network maps of the number of households moving from each service to another in the optimal allocation

between services. It is clear that the main mechanism of improvement is a flow where a significant number of households are being placed in transitional housing rather than shelters; in order to make room for these, households move from transitional housing to prevention, and from prevention to shelters. This flow indicates a potentially complex mechanism for improving outcomes, since it is not simply a two-way swap between services.

We explore further who benefits in optimization to assess potential inequities. We build random forest models using the default hyperparameter values listed for the classification problem of predicting whether a household has a higher or lower probability of reentry after optimal allocation. We chose random forests due to the ease of producing measures of variable importance from a random forest model. The models have access to the entire original set of features, but ignore service type. The relative importance of each feature for prediction (calculated using the mean decrease in accuracy of features – a standard permutation test used in random forest feature importance) provides insights into the key characteristics that differentiate those who improve or worsen their reentry probability. The out-of-bag error for the random forest model was 0.09 and the AUC was 0.96. Figure 3 plots the 30 most influential variables. Some of the most important features are housing status at entry, the number of hotline calls prior to entry, and prior residence.

Perhaps the most striking discovery to emerge from the analysis is that the optimal allocation seems to help those who stand out as being *more* in need. Households benefited most by reallocation disproportionately are homeless upon entry and make frequent calls to the hotline for help; they also are more likely to reside in non-federally funded homeless services (primarily provided through local religious organizations), substance abuse treatment facilities, or with family. Moreover, reallocation benefits households more likely to report a disability who wait longer for entry into services. Households harmed by optimization, on the other hand, are more likely to be at imminent risk or stably housed upon entry, first time hotline callers with briefer waits for services, and in their own or rental units; household heads also are somewhat older and more likely to have children. The no change group also

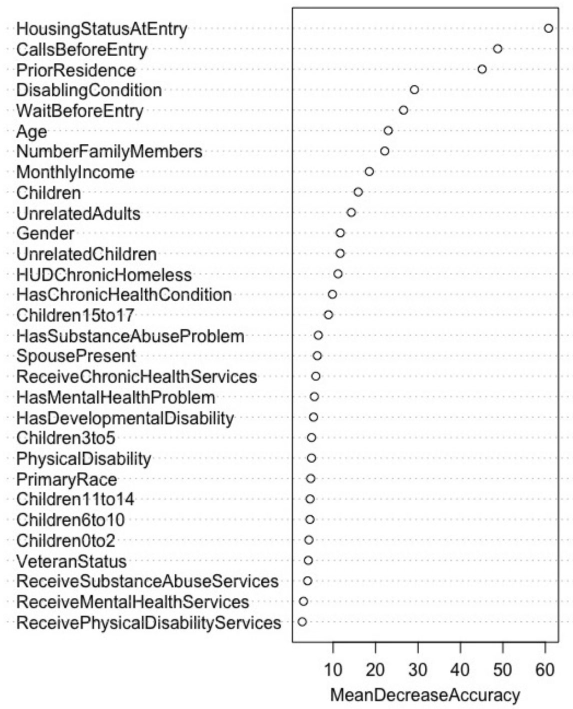


Figure 3: Plot of the mean decrease in accuracy of features for predicting whether the optimal allocation will increase or decrease a household's probability of reentry

experience stable housing in their own units upon entry. Table 7 summarizes comparisons of household characteristics by reallocation outcomes. All differences in continuous variables between the group who improved versus harmed were tested using a Student's t-test and are significant below the $p < .001$ level.

Overall, these results suggest an ability to improve upon the allocation rules used by the homeless system. To note, although more than one optimal solutions could exist, we find evidence only for a single solution across runs. Interestingly, the efficiency gains are achieved primarily through "shuffling" households between emergency shelters (which is a uniformly poor service), prevention (which may be appropriate for more vulnerable households than previously believed), and transitional housing (an intense and expensive service with higher efficacy). There is clearly some household-level heterogeneity that could potentially be exploited to achieve gains.

5.3 Constraining Increased Probability of Reentry

Another important dimension of fairness raised in algorithmic decision-making pertains to the local costs of redistributing resources. Inefficiencies in the original allocation may be because decision-makers are prioritizing equity by assigning more vulnerable households to more intensive services (whether the measurement of vulnerability corresponds to the actual notion we care about is a separate question)(Fowler et al., 2019a). Of course, this

Feature	Improved Group (n = 3967) %/Mean (SD)	Harmed Group (n = 2825) %/Mean (SD)	No Change Group (n = 3251) %/Mean (SD)	Total (n = 10043) %/Mean
Housing Status At Entry: Homeless	89.42	4.18	6.41	10.72
Housing Status At Entry: At imminent risk of losing housing	5.35	50.17	44.48	34.63
Housing Status At Entry: At-risk of homelessness	4.43	45.87	49.70	11.46
Housing Status At Entry: Stably Housed	12.54	40.07	47.39	2.86
Housing Status At Entry: Client doesn't know	67.43	9.68	22.89	40.33
Calls Before Entry	4.15 (6.60)	0.94 (2.55)	2.51 (5.31)	2.72 (5.46)
Prior Residence: Emergency Shelter	71.11	8.75	20.14	12.85
Prior Residence: Transitional housing for homeless persons	65.50	23.08	22.50	5.97
Prior Residence: Substance abuse treatment facility or detox center	70.52	7.70	21.78	16.72
Prior Residence: Staying or living in with family member	54.61	18.60	26.80	11.78
Prior Residence: Rental by client no ongoing housing subsidy	4.93	44.87	50.20	24.45
Prior Residence: Owned by client no ongoing housing subsidy	3.94	57.03	39.03	12.88
Head of Household Has Disabling Condition: No	37.62	28.34	34.05	82.82
Head of Household Has Disabling Condition: Yes	50.00	25.63	24.47	15.03
Head of Household Has Disabling Condition: Don't Know	38.50	38.03	23.47	2.12
Head of Household Has Disabling Condition: Refused	50.00	0.00	50.00	0.02
Wait Before Entry	277.19 (476.18)	115.51 (354.23)	266.04 (476.22)	228.10 (450.80)
Age of Head of Household	38.39 (13.22)	40.87 (12.59)	39.75 (12.24)	39.53 (12.77)
Number of Family Members	1.57 (1.15)	2.57 (1.56)	1.96 (1.29)	1.98 (1.38)
Monthly Income	892.58 (1069.03)	1906.01 (2091.89)	1524.80 (2066.83)	1382.31 (1800.82)
Number of Children	0.54 (1.10)	1.34 (1.43)	0.88 (1.23)	0.88 (1.28)

Table 7: Summary statistics for the most influential features for determining which households will benefit from the optimal allocation (due to the large number of prior residence categories, those making up less than 5% of the population were omitted from the table)

idea may be flawed in that some of these "more vulnerable" households may actually be equally well-served by less intensive services.

One way to potentially deal with fairness concerns like these is to make them explicit in the optimization. As an example, we consider what happens if we add a constraint that prevents any household from suffering too high a predicted cost, in terms of predicted increases in probability of reentry, from the change in allocation. For example:

$$\sum_j p_{ij} x_{ij} \leq \sum_j p_{ij} y_{ij} + \delta \; \forall i$$

where each y_{ij} is a binary variable representing whether or not household i was originally placed in service j. And δ is a constraint which keeps households from being allocated to a service in which their predicted probability of reentry is more than δ percentage points higher than that of the service they participated in originally.

To illustrate the results of the allocation when this constraint is added, Figure 4 shows the distribution of changes in the the new allocation when δ is set to 5 percentage points. The hard threshold of course prevents any negative changes of greater than 5 percentage points. When we include this constraint, the solution to the optimization problem yields an allocation with a predicted 2238 households (22.29%) reentering the system within two years. This is obviously higher than the optimized allocation without the constraint, but still a 19.06% decrease compared to the predicted reentry number for the original allocation. Looking again at individual households, 2776 (27.64%) are allocated into a service that lowers probability of reentry, 5822 (57.97%) are allocated into the original assignment, and 1445 (14.39%) are allocated into a service that increases probability of reentry. The majority of households who do worse suffer very small penalties.

We empirically investigate the influence of imposing more and less restrictive fairness constraints on reentry rates. Figure 5 shows the percentage reduction in expected reentries

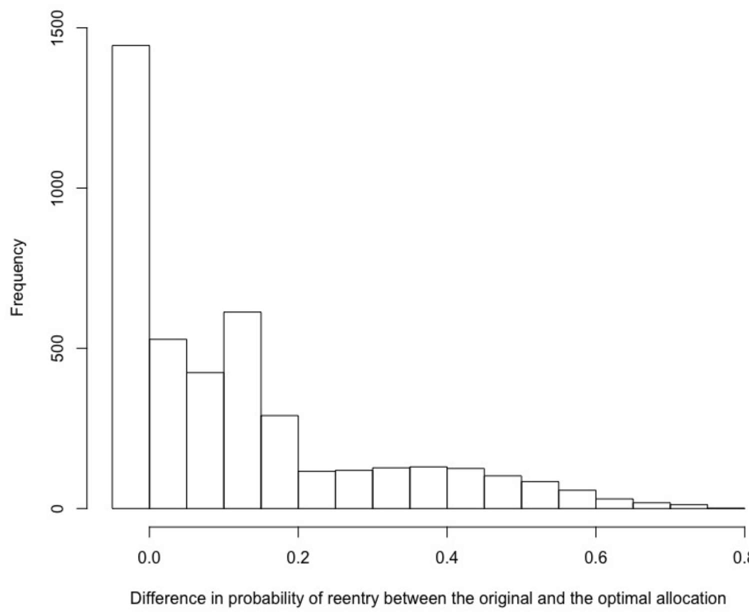


Figure 4: Histogram of improvement in reentry probability under the constrained optimized allocation (the 5822 individuals whose probability of reentry was unchanged are not included)

as a function of δ (how much each household's predicted reentry probability is allowed to increase in the optimal allocation). That is, how much predicted cost households are allowed to incur from the change in allocation. An inflection point exits between four and five percentage point constraints, such that gains in reentry reductions begin to level off. Depending on fairness considerations, thresholds could be set higher to maximize reductions or lower to promote equity for more vulnerable households. A particularly interesting result from this investigation is that, even when the constraint does not allow any household's predicted reentry probability to increase at all, we still achieve a 14.31% reduction in expected reentries. Therefore, we can produce gains in efficiency even in the presence of strict fairness constraints.

Discussion

Our work tests the feasibility of using data-driven counterfactual approaches to inform policies that guide homeless service provision. We analyze the potential for different allocation mechanisms to improve outcomes using counterfactual estimates of probability of reentry into the system. Our results suggest that optimal weekly assignments reduces system reentries substantially. However, optimization of system-wide service delivery withholds useful services for one-third of households. Although the average harm to households pails in comparison to the benefits for other households, the results emphasize that optimal reallocation of services fails to improve the outcomes of all households in the homeless system. As-

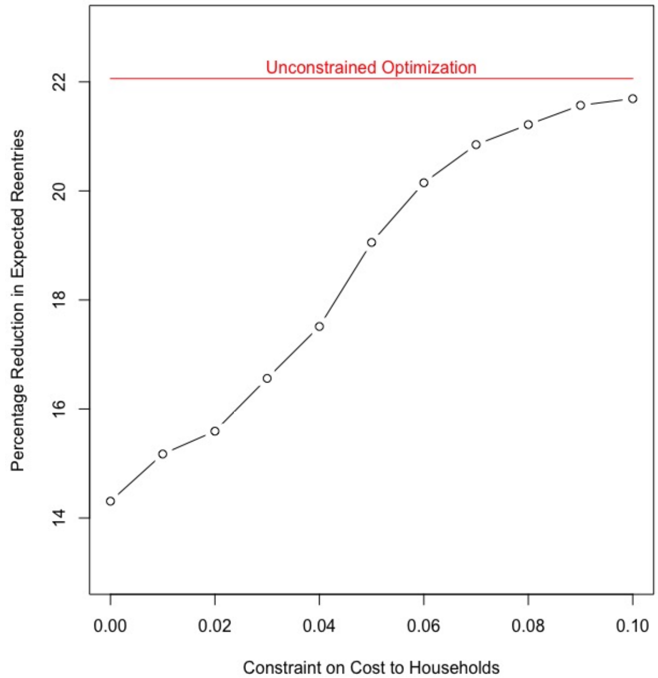


Figure 5: Graph showing percent decrease in expected number of reentries as a function of constraint on how much a household's predicted reentry probability is allowed increase in the optimal allocation

suming the original allocation to be fair, models explore the imposition of an approximate fairness constraint that avoid households from being reallocated to services that worsen the probability of reentry into the system compared to the original allocation. Results show smaller but meaningful reductions in reentries into the homeless system using fair data-driven allocations of services

Our findings demonstrate the critical importance of fairness and justice considerations in the design of algorithmic allocations of homeless services delivery. The assumptions, implications, and potential unintended consequences must be thoroughly analyzed and addressed before implementing data-driven decision-making. One potential solution allows workers to override certain allocation decisions. The idea has previously been adopted as part of a homelessness prevention screening instrument used in New York City (Shinn et al., 2013). Shinn and colleagues note that analysis of the reasons behind these overrides can help to inform future models of this type. The addition of potential override reasons to an allocation model could help to increase fairness and inform re-calibrations of models. It also makes the transition to an allocation program smoother by allowing homeless service workers to maintain control over allocations.

The results presented here must be considered in the context of limitations of this kind of study. It is difficult to rule out all potential confounds for treatment estimates. Our models leverage all available data from homeless services for predictions, and extensive sensitivity analyses provide some confidence in the results. However, the observational nature of the data constrains modeling for variables we were not aware of or to which we did not have access. If the estimated treatment effects are biased, this would inherently worsen efficiency gains by introducing unreliability.

Another key limitation concerns the potential for unobserved inequities in homeless service delivery. Administrative records only collect information on services provided; models remain vulnerable to service decisions that intentionally (i.e., explicit bias) or unintentionally (i.e., implicit bias) disadvantage specific groups. As illustrated in prior applications, algorithmic decision making risks perpetuating systematic inequalities captured in the data (Ensign et al., 2018; Pleiss et al., 2017; Corbett-Davies et al., 2017). Surprisingly, initial tests in the present study suggest optimal allocation disproportionately advantages more vulnerable households. The unexpected findings potentially reveal counterproductive assumptions guiding service delivery. Currently, homeless policies prioritize scarce intensive services for more vulnerable households, whereas the data-driven allocation maximizes timely receipt of preventive services for first time entries into the homeless system (United States Congress, 2009). These findings are consistent with a growing body of evidence on community-wide benefits of homelessness prevention (Fowler et al., 2019b). Insights from the present study introduce new avenues for future work that informs data-driven homeless service delivery. Further investigation into heterogeneous effects of different homeless services offers opportunities to ask key policy questions of what works for whom. This is especially true for prevention services that unexpectedly show promise at first time entry. In addition, deeper investigation into winners and losers of data-driven allocation needs to test for potential disparities. Fairness considerations must extend to assess whether specific groups are being disproportionately reassigned to certain services (e.g. shelter versus prevention). Answering questions like this would help us learn how to decrease the number of households harmed by this type of service allocation.

In sum, our study demonstrates both the potential of, and the need for caution in, data-driven homeless service delivery. Although machine learning improves efficiency, fairness considerations arise that require careful implementation in practice. Data-driven insights also raise questions regarding policies that underlie service delivery – fitting an algorithms-in-the-loop process (Green & Chen, 2019). This study opens new lines of inquiry for designing and testing computational approaches that promote social good.

Acknowledgments

A preliminary version of this work appeared in the Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI 2019) (Kube, Das, & Fowler, 2019). We are grateful for support from NSF awards 1527037, 1560191, 1910392, and 1927422, as well as Administration for Children and Families Childrens Bureau Award 90CA1815. We appreciate the assistance of Municipal Information Systems Inc. (MISI) for coordinating data access, as well as the households who it represents. We are grateful to numerous local homeless providers, who graciously contributed insights throughout the study. We also benefited from comments on this work from Dennis Culhane, John P. Dickerson, David Savitt, Phebe

Vayanos, participants in the CAIS Seminar at the University of Southern California, and the Sawyer Seminar at Johns Hopkins University.

References

Abdulkadiroglu, A., Pathak, P. A., Roth, A. E., & Sonmez, T. (2005). The Boston public school match. *American Economic Review*, 368–371.

Akbarpour, M., Li, S., & Gharan, S. O. (2020). Thickness and information in dynamic matching markets. *Journal of Political Economy*, *128*(3), 783–815.

Anshelevich, E., Chhabra, M., Das, S., & Gerrior, M. (2013). On the social welfare of mechanisms for repeated batch matching. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 60–66.

Anshelevich, E., & Das, S. (2010). Matching, Cardinal Utility, and Social Welfare. *ACM SIGECom Exchanges*, *9*(1).

Anshelevich, E., Das, S., & Naamad, Y. (2013). Anarchy, stability, and utopia: Creating better matchings. *Autonomous Agents and Multi-Agent Systems*, *26*(1), 120–140.

Azizi, M. J., Vayanos, P., Wilder, B., Rice, E., & Tambe, M. (2018). Designing fair, efficient, and interpretable policies for prioritizing homeless youth for housing resources. In *International Conference on the Integration of Constraint Programming, Artificial Intelligence, and Operations Research*, pp. 35–51. Springer.

Bansak, K., Ferwerda, J., Hainmueller, J., Dillon, A., Hangartner, D., Lawrence, D., & Weinstein, J. (2018). Improving refugee integration through data-driven algorithmic assignment. *Science*, *359*(6373), 325–329.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32.

Brown, A., Chouldechova, A., Putnam-Hornstein, E., Tobin, A., & Vaithianathan, R. (2019). Toward algorithmic accountability in public services: A qualitative study of affected community perspectives on algorithmic decision-making in child welfare services. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 41:1–12. ACM.

Brown, M., Cummings, C., Lyons, J., Carrión, A., & Watson, D. P. (2018). Reliability and validity of the Vulnerability Index-Service Prioritization Decision Assistance Tool (VI-SPDAT) in real-world implementation. *Journal of Social Distress and the Homeless*, *27*(2), 110–117.

Butaru, F., Chen, Q., Clark, B. J., Das, S., Lo, A. W., & Siddique, A. R. (2016). Risk and risk management in the credit card industry. *Journal of Banking and Finance*, *72*, 218–239.

Chen, C., Zheng, L., Srinivasan, V., Thomo, A., Wu, K., & Sukow, A. (2016). Conflict-aware weighted bipartite b-matching and its application to e-commerce. *IEEE Transactions on Knowledge and Data Engineering*, *28*(6), 1475–1488.

Chipman, H. A., George, E. I., & McCulloch, R. E. (2007). Bayesian ensemble learning. *Advances in Neural Information Processing Systems*, *19*, 265–272.

Chipman, H. A., George, E. I., McCulloch, R. E., et al. (2010). BART: Bayesian additive regression trees. *The Annals of Applied Statistics*, *4*(1), 266–298.

Chouldechova, A., Benavides-Prado, D., Fialko, O., & Vaithianathan, R. (2018). A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. In *Proceedings of the Conference on Fairness, Accountability and Transparency*, pp. 134–148.

Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (2017). Algorithmic decision making and the cost of fairness. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 797–806. ACM.

Culhane, D. P., Park, J. M., & Metraux, S. (2011). The patterns and costs of services use among homeless families. *Journal of Community Psychology*, *39*(7), 815–825.

Dickerson, J. P., Manlove, D. F., Plaut, B., Sandholm, T., & Trimble, J. (2016). Position-indexed formulations for kidney exchange. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pp. 25–42. ACM.

Dilkina, B., & Gomes, C. P. (2010). Solving connected subgraph problems in wildlife conservation. In *International Conference on Integration of Artificial Intelligence (AI) and Operations Research (OR) Techniques in Constraint Programming*, pp. 102–116. Springer.

Ensign, D., Friedler, S. A., Neville, S., Scheidegger, C., & Venkatasubramanian, S. (2018). Runaway feedback loops in predictive policing. In *Conference on Fairness, Accountability and Transparency*, pp. 160–171. PMLR.

Fowler, P. J., Hovmand, P. S., Marcal, K. E., & Das, S. (2019a). Solving homelessness from a complex systems perspective: Insights for prevention responses. *Annual Review of Public Health*, *40*(1), 465–486. PMID: 30601718.

Fowler, P. J., Wright, K., Marcal, K. E., Ballard, E., & Hovmand, P. S. (2019b). Capability traps impeding homeless services: A community-based system dynamics evaluation. *Journal of Social Service Research*, *45*(3), 348–359.

Green, B., & Chen, Y. (2019). Disparate Interactions: An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments. In *Proceedings of the Conference on Fairness, Accountability, and Transparency - FAT* '19*, pp. 90–99, Atlanta, GA, USA. ACM Press.

Gubits, D., Shinn, M., Wood, M., Brown, S. R., Dastrup, S. R., & Bell, S. H. (2018). What Interventions Work Best for Families Who Experience Homelessness? Impact Estimates from the Family Options Study. *Journal of Policy Analysis and Management*, *37*(4), 835–866.

Henry, M., Mahathey, A., Morrill, T., Robinson, A., Shivji, A., & Watt, R. (2018). The 2018 Annual Homeless Assessment Report (AHAR) to Congress: Part 1: Point-in-Time Estimates of Homelessness. Tech. rep., United States Department of Housing and Urban Development, Washington D.C.

Hill, J. L. (2011). Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, *20*(1), 217–240.

HUD (2012). Homeless emergency assistance and rapid transition to housing: Continuum of care program; interim final rule. *Federal Register*, *77*(147), 45422–45467.

Johansson, F., Shalit, U., & Sontag, D. (2016). Learning representations for counterfactual inference. In *International Conference on Machine Learning*, pp. 3020–3029.

Khadduri, J., Leopold, J., Sokol, B., & Spellman, B. (2010). *Costs Associated with First-Time Homelessness for Families and Individuals*. Dept of Housing and Urban Development.

Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction policy problems. *The American Economic Review*, *105*(5), 491–495.

Kominers, S. D., Teytelboym, A., & Crawford, V. P. (2017). An invitation to market design. *Oxford Review of Economic Policy*, *33*(4), 541–571.

Kube, A., Das, S., & Fowler, P. J. (2019). Allocating interventions based on predicted outcomes: A case study on homelessness services. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*, 622–629.

Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, *2*(1-2), 83–97.

Li, Z., Lieberman, K., Macke, W., Carrillo, S., Ho, C.-J., Wellen, J., & Das, S. (2019). Incorporating compatible pairs in kidney exchange: A dynamic weighted matching model. In *Proceedings of the ACM Conference on Economics and Computation*, pp. 349–367.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group, New York, NY, USA.

Pleiss, G., Raghavan, M., Wu, F., Kleinberg, J., & Weinberger, K. Q. (2017). On fairness and calibration. In *Advances in Neural Information Processing Systems*, pp. 5684–5693.

Roth, A. E., Sonmez, T., & Utku Unver, M. (2005). Pairwise kidney exchange. *Journal of Economic Theory*, *125*(2), 151–188.

Roth, A. E., & Sotomayor, M. (1990). *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Econometric Society Monograph Series. Cambridge University Press, Cambridge UK.

Shinn, M., Brown, S. R., Spellman, B. E., Wood, M., Gubits, D., & Khadduri, J. (2017). Mismatch Between Homeless Families and the Homelessness Service System. *Cityscape (Washington, D.C.)*, *19*(3), 293–307.

Shinn, M., Greer, A. L., Bainbridge, J., Kwon, J., & Zuiderveen, S. (2013). Efficient targeting of homelessness prevention services for families. *American Journal of Public Health*, *103*(S2), S324–S330.

Trapp, A. C., Teytelboym, A., Martinello, A., Andersson, T., & Ahani, N. (2021). Placement optimization in refugee resettlement. *Operations Research*. To appear.

United States Congress (2009). Homeless emergency assistance and rapid transition to housing act..

Weinberger, D. (2018). Don't make artificial intelligence artificially stupid in the name of transparency. *Wired*.

Yadav, A., Chan, H., Xin Jiang, A., Xu, H., Rice, E., & Tambe, M. (2016). Using social networks to aid homeless shelters: Dynamic influence maximization under uncertainty. In *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*, pp. 740–748.

Appendix A. Descriptive Statistics

Table 8 provides descriptive statistics for the dataset.

Feature	**Total** n=13940
	% / M(SD)
Household Characteristics	
Number of Household Members	1.94(1.36)
Spouse Present	6.55
Number of Children	0.84(1.26)
Number of Children Ages 0 to 2	0.16(0.43)
Number of Children Ages 3 to 5	0.14(0.41)
Number of Children Ages 6 to 10	0.20(0.53)
Number of Children Ages 11 to 14	0.15(0.43)
Number of Children Ages 15 to 17	0.10(0.34)
Number of Unrelated Adults	0.04(0.22)
Number of Unrelated Children	0.06(0.36)
Number of Calls Before Entry	2.82(5.34)
Wait Before Entry (in days)	192.70(402.18)
Monthly Income (in US Dollars)	1328.73(1792.28)
Head of Household Characteristics	
Female	66.19
Age (years)	40.04(12.66)
White	15.22
African American	83.24
Hispanic or Latino Ethnicity	1.17
Veteran	4.71
Disabling Condition	12.40
Physical Disability	14.05
Received Physical Disability Services	5.55
Developmental Disability	1.92
Received Developmental Disability Services	0.36
Chronic Health Condition	25.07
Received Chronic Health Services	14.19
HIV/AIDS	0.44
Received HIV/AIDS Services	0.19
Mental Health Problem	21.56
Received Mental Health Services	8.67
Alcohol Abuse Problem	4.19
Drug Abuse Problem	9.71
Both Alcohol and Drug Abuse Problem	5.83
Received Substance Abuse Services	9.66
Domestic Violence Survivor	0.59
Chronically Homeless	2.14
Homeless	8.29
At Imminent Risk of Losing Housing	25.47
At Risk of Homelessness	8.91
Stably Housed	2.88
Coming from Emergency Shelter	12.08
Coming from Transitional Housing	6.18
Coming from Substance Abuse Treatment Facility or Detox Center	8.06
Coming from Hospital or other Residential Non-psychiatric Medical Facility	1.37
Coming from a Family Member's Residence	11.61
Coming from a Friend's Residence	3.19
Coming from a Place not Meant for Habitation	4.73
Coming from a Rental with Housing Subsidy	1.42
Coming from a Rental without Housing Subsidy	17.80
Coming from a Residence Owned by Client without Housing Subsidy	19.61

Table 8: Summary of the dataset by service type