Quadratic Optimization Models for Balancing Preferential Access and Fairness: Formulations and Optimality Conditions

Christian Schmitt
Department of Data Science, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Bavaria 91058, Germany

Bismark Singh
School of Mathematical Sciences, University of Southampton, Southampton, S017 1BJ, UK
b.singh@southampton.ac.uk

Typically, within facility location problems, fairness is defined in terms of accessibility of users. However, for facilities perceived as undesirable by communities hosting them, fairness between the usage of facilities becomes especially important. Limited research exists on this notion of fairness. To close this gap, we develop a series of optimization models for the allocation of populations of users to facilities such that access for users is balanced with a fair utilization of facilities. The optimality conditions of the underlying non-convex quadratic models provide a precise tradeoff between accessibility and fairness. We define new classes of fairness, and a metric to quantify the extent to which fairness is achieved in both optimal and suboptimal allocations. We show a continuous relaxation of our central model is sufficient to achieve a perfect extent of fairness, while a special case reduces to the classical notion of proportional fairness. Our work is motivated by pervasive ecological challenges faced by the waste management community as policymakers seek to reduce the number of recycling centers in the last few years. As a computational case study, applying our models on data for the state of Bavaria in Germany, we find that even after the closure of a moderate number of recycling centers, large degrees of access can be ensured provided the closures are conducted optimally. Fairness, however, is impacted more, with facilities in rural regions shouldering larger loads of visiting populations than those in urban regions.

Key words: Quadratic combinatorial optimization; KKT optimality conditions; Fairness; Waste management; Facility location problems

History:

1. Introduction

We revisit the class of Facility Location Problems (FLPs), which have a rich history in the optimization literature. Specifically, we are interested in so-called *undesirable* or *semi-desirable* facilities; i.e., facilities that exert a negative impact on the surrounding community, while still providing a necessary service. Examples of such facilities include sanitary landfills, airports, or thermal stations (Erkut and Neuman 1989). Users prefer certain facilities, while facilities have finite capacities. There are two aims we pursue: (i) select a subset of facilities for the allocation, and (ii) provide an assignment of users to the selected facilities. We determine the allocation and assignment in a manner that is both *fair* to the community surrounding a facility and *accessible* to the users seeking this facility; we make the two terms—"fair" and "accessible"—precise later in this work. A high utilization of a facility, that results from assigning a large amount of users to it, can have significant detrimental impact on the community that hosts this facility. Thus, an assignment that is highly accessible for the users could disproportionately burden some communities. Analogously, an allocation that is highly fair for the community around a facility could result in assignments that are inaccessible, or biased, to the users. We formulate optimization models where optimal solutions balance these two goals, and the aim of this work is to *analytically* (where possible) quantify such compromises. To this end, we present general results—using the optimality conditions of our models—that determine this tradeoff relationship between accessibility and fairness.

Preferences of users to facilities in FLPs are typically modeled via weights that quantify accessibility, e.g., distances between the user and the facility. In locating desirable facilities, a policymaker might seek to reduce these distances. However, for locating undesirable facilities, a policymaker needs to incorporate additional criteria based on fairness to the communities hosting these facilities (henceforth, fairness among the facilities). There is an extensive body of literature devoted to addressing such fairness concerns in this "obnoxious" FLP. One typical approach is to maximize the distance of users to facilities in some fashion (Cappanera 1999). A second well-studied approach is to minimize the population lying within a given radius of a facility (Plastria and Carrizosa 1999). Inequity in assignments is also long acknowledged in the resource allocation literature; thus, various measures of equity and fairness have been proposed, see, e.g., Marsh and Schilling (1994). Bertsimas et al. introduce several classes of the optimization problem of fairly allocating resources, where recipients have varying utilities (Bertsimas et al. 2011). Fairness criteria play an especially important role in communication networks where aspects of this problem are studied, see, e.g., Kelly et al. (1998). These lead to competing definitions of fairness, and two well-studied notions are those of *max-min* fairness and *proportional* fairness (Bertsimas et al. 2011, Kaplan 1974, Kelly et al. 1998, Singh 2020). We revisit these concepts later in this work.

However, only a few works consider fairness in the sense of disparities in usage of facilities. Two exceptions are Berman et al. (2009) and Marín (2011), where the authors seek to achieve fairness by minimizing the maximum load a facility carries and by balancing the maximum and minimum number of users assigned to a facility, respectively. In a similar spirit as these two works, our notion of fairness among facilities dwells from reducing disparities in *utilization* of facilities, while ensuring large degrees of access for the users visiting these facilities. However, the two goals of maximizing accessibility of users to a facility and minimizing disparities in the utilization of facilities are conflicting. Two factors cause further challenges: (i) users have varying preferences to certain facilities, such as shorter travel times or preferences that depend on individual choices (Tversky and Simonson 1993), and (ii) the preferred facility of a user might not have enough capacity to accommodate it (Pirkul and Schilling 1988). With this background, there are two broad components to this work.

The first component is a theoretical contribution to the rich class of FLPs. We begin by proposing a quadratic-binary optimization model that is uniquely distinguished from traditional FLPs by the choice of its objective function; in Section 2, we discuss in detail the reasons for this modeling choice. Our model is induced by a new notion of fairness that encompasses both of the competing goals we mention above. Feasible solutions of this model that are suboptimal pay a price both in terms of the maximum access of users and the minimum dispersion in the utilization of the facilities. Interestingly, even solutions that are optimal achieve fairness only to a limited extent. With this motivation, we construct a metric that quantifies the *extent* of fairness. The natural follow-up question is when—if at all—is a perfect extent of fairness achieved. We prove that a suitably relaxed version of our optimization model achieves this in an optimal solution, thereby providing one sufficient condition. Next, we show that our notion of fairness generalizes existing results on proportional fairness; specifically, we prove that a special case of this relaxed model achieves proportional fairness.

The second component of this work is inspired by a pervasive problem in the waste management community—positioning of recycling centers. As of 2018, Germany is among the leading countries worldwide in the proportion of waste that is recycled (Kaza et al. 2018). In Germany, the "Wertstoffhöfe" (or, recycling centers) are facilities where the public is required to dispose of waste that is not regularly collected from households. A few examples of such waste include construction waste, recyclable electronics, large appliances, and scrap metal. However such recycling locations are perceived as undesirable facilities, and these negative perceptions are further enhanced when a region is forced to shoulder a disproportionately large burden of these facilities (Morell 1984). These concerns are justified as recycling centers pose health and environmental hazards, e.g., recycling centers have been found to increase risks of fires (Ibrahim 2020), and storing electronic waste can lead to health risks in the surrounding community through trace metal pollution in soil and road dust (Yekeen et al. 2016). As a result, the number of such recycling centers has steadily decreased within the state of Bavaria during the last few years (Bayerisches Landesamt für Umwelt 2015); the number of recycling centers in the years 2016, 2017, 2018, and 2019 were 1,624, 1,597, 1,583, and 1,578, respectively (Bayerisches Landesamt für Umwelt 2020). This leads to a natural question that policymakers face: what is an optimal selection of a limited number of recycling centers to keep open? Given this budget of recycling centers, an efficient usage of the open facilities by the public is paramount for a state's recycling policy. We apply our optimization models as a case study for the entire set of users and recycling locations of Bavaria and provide extensive computational experiments. Here, our analysis provides data-driven evidence on significant disparity between rural and urban regions both in the treatment of recycling centers and in the populations visiting them.

The structure of the rest of this article is as follows. In Section 2, we propose our central and relaxed optimization models plus provide our definitions and metrics for fairness. In Section 3, we demonstrate the connection of our models to existing results on proportional fairness. Section 4 summarizes the estimation of data from Bavaria for our computational case study in Section 5. We provide a concluding discussion in Section 6, and further details, analysis, and proofs in the appendices.

2. Mathematical Models

Consider a set of users $i \in I$ with populations $U_i > 0$ and a set of facilities $j \in J$ with capacities $C_j > 0$. Populations of users have a preference to facilities measured in terms of weights or probabilities; without loss of generality we assume $0 < P_{ij} \leq 1$. For example, populations might prefer facilities that are closer to their residences or those that are open on weekends with a higher probability than others. In the absence of this preference, if user i is allocated to facility j then its entire population U_i visits j. Instead, the preference discounts the population of i that actually visits j to $U_i P_{i,j} < U_i$. Then, the following optimization model describes our central problem of assigning users to facilities in a fair manner:

$$z^* = \min_{x,y} \sum_{j\in J} C_j\Big(1 - \frac{\sum_{i\in I} U_i P_{i,j} x_{i,j}}{C_j}\Big)^2 \tag{1a}$$

$$\text{s.t.} \sum_{i\in I} U_i P_{i,j} x_{i,j} \leq C_j \qquad \forall j \in J : \alpha_j \tag{1b}$$

$$\sum_{j\in J} y_j \leq B \qquad : \beta \tag{1c}$$

$$\sum_{j\in J} x_{i,j} = 1 \qquad \forall i \in I : \nu_i \tag{1d}$$

$$y_j \geq x_{i,j} \qquad \forall i \in I, j \in J : \mu_{i,j} \tag{1e}$$

$$x_{i,j} \in \{0,1\} \qquad \forall i \in I, j \in J : \gamma_{i,j} \tag{1f}$$

$$y_j \in \{0,1\} \qquad \forall j \in J : \delta_j \tag{1g}$$

The optimization model (1) is a quadratic-binary program. Here, $\alpha_j, \beta, \nu_i, \mu_{i,j}, \gamma_{i,j}, \delta_j$, denote the dual variables for equations (1b)-(1g), respectively; we revisit these later. Note that the binary restrictions in constraint (1g) can be written as $y_j(1-y_j)=0$ with the dual variables interpreted accordingly. Model (1) is non-convex due to the binary restrictions on the decision variables, however its relaxation is a convex quadratic program. To understand model (1), we begin by defining two key quantities below.

$$a_j = \textstyle\sum_{i\in I} U_i P_{i,j} x_{i,j},\ \forall j \in J, \tag{2a}$$

$$u_j = \frac{\sum_{i\in I} U_i P_{i,j} x_{i,j}}{C_j},\ \forall j \in J. \tag{2b}$$

First, in equation (2a) the quantity a_j denotes the total population assigned to facility j; this quantity is a measure of the *access* of facility j for all the users. Second, in equation (2b) the quantity u_j denotes the *utilization* of facility j; i.e., the fraction of its total capacity that is actually being used. Conversely, the idleness of facility j is given by $1-u_j$. Then, model (1) minimizes the capacity-weighted sum of squared idleness; we explain this objective function in detail below. Constraint (1b) ensures the utilizations are no more than 1, while constraint (1c) allows no more than a budget of B facilities to open. The binary variables $x_{i,j}$ and y_j govern whether user i is assigned to facility j and whether facility j is open, respectively; this is ensured by constraints (1f) and (1g), respectively. Constraint (1d) ensures a user i is assigned to only one facility; i.e., the entire population of user i—discounted by the preferences—is allocated only to a single facility. Thus, the fraction $\sum_{j\in J} U_i(1-P_{i,j})x_{i,j}$

of the population of user i is left unallocated to any facility as an artifact of user i being stringent in its preferences. Constraint (1e) enforces facility j is open if any user is assigned to it, while if j is closed then no user is assigned to it; i.e., $y_j = 1$ if $x_{i,j} = 1$ for any i, while if $y_j = 0$, then $x_{i,j} = 0, \forall i \in I$.

In model (1), constraints (1b), (1d)-(1g) are similar to those of a traditional capacitated FLP (see, e.g., (Wolsey 1998)) with two additional restrictions. The first is a preferential access to facilities—this preference is reflected via the coefficients $U_i P_{i,j}$. The second is a budget on the number of facilities—this is enforced via constraint (1c). Proposition A.1 in Appendix A.1, motivated by the capacitated FLP, provides a valid inequality for model (1). Analogously, constraints (1c)-(1g) are similar to those of a p-median problem (see, e.g., (ReVelle and Swain 1970)). Then, the binary restrictions in (1g) can be replaced with their continuous relaxation, $0 \le y \le 1$; since there exists an optimal solution where y is binary if x is binary (see, Proposition A.2 in Appendix A.1). However, the binary restrictions in equation (1f) cannot be replaced with their continuous relaxation; this is unlike the traditional p-median problem due to the choice of our objective function (see, Proposition A.3 in Appendix A.1 for a counterexample).

An important feature distinguishing model (1) from traditional FLPs is our choice of the objective function (1a). Our choice is central to the discussion and results in this section, and the rationale behind this choice is subtle. We first provide an intuitive explanation and then lay out the formal reasoning that substantiates our choice. Assigning each user to its most accessible facility, even when the capacities allow, results in a disproportionately large degree of utilization for the highly-accessible facilities. On the other hand, allocating facilities to a capacity-proportional amount of utilization might lead to inaccessible assignments for remote users. Thus, as we mention in Section 1, there are two conflicting aims that we pursue and the objective function (1a) precisely captures both of these. First, we seek to achieve fairness between facilities by ensuring that no facility carries a disproportionate burden of assigned users; the proportion of burden is determined by the facility's capacity. Simultaneously, we seek to maximize users' access to the facilities. The terms in the objective function (1a) get smaller for a larger utilization u_j, and, in turn, from a larger access to the facility j. The sum of the squared idleness seeks to reduce the disparity between the utilization of the different facilities; the weights C_j determine the relative weighting of the utilization of the facilities. To this end, model (1) provides a tradeoff between access and fairness. With this background, we are now ready to define our notion of fairness among facilities that makes this tradeoff precise.

DEFINITION 1 (FAIRNESS). Consider an allocation with utilization defined by equation (2b), and two distinct open facilities $j, j' \in J$.

(i) The allocation to the ordered pair (j, j') is said to have a *warranted fairness* if $P_{i,j} \geq P_{i,j'}$ for every user $i \in I$ assigned to j.

(ii) The allocation to the ordered pair (j, j') is said to have a *compensatory fairness* if there exists a user $i \in I$ assigned to j with $P_{i,j} < P_{i,j'}$ and $u_j \leq u_{j'}$.

Then, an allocation is said to be *fair* if it has a warranted or compensatory fairness for all ordered pairs of distinct open facilities (j, j'). □

Definition 1 classifies a fair solution between two facilities as follows. First, fairness is only defined for facilities that are open $(y > 0)$. In a choice between two open facilities, if all users are assigned to their preferred facility then there is no basis for a competition. This denotes a *warranted fairness*. Such an allocation is most favorable for a policymaker as it allows no grounds for unfairness. However, now consider a user i that prefers an open facility j' over another open facility j, yet in a feasible solution i is assigned to j. Such a solution is fair for the respective facilities only if the assignment seeks to balance out the utilization of the facilities; i.e., j is utilized at most as much as j'. Such an allocation, although not completely fair in terms of preferences towards facilities, is compensated for fairness by a larger utilization of the preferred facility. Then, this allocation has a *compensatory fairness* according to Definition 1 if there is at least one user assigned in this fashion. Pairs of open facilities with neither warranted nor compensatory fairness are allocated unfairly.

A natural follow-up question is the extent to which fairness is achieved in optimal solutions; i.e., the proportion of open pairs allocated fairly. The following definition quantifies this.

DEFINITION 2 (DEGREE OF FAIRNESS). Let $J_O = \{j \in J : y_j = 1\}$ denote the set of open facilities and let $I_j = \{i \in I : x_{i,j} = 1\} \subseteq I, \forall j \in J_O$ denote the set of users that are assigned to an open facility j. Further let $k_{j,j'} = \mathbb{1}(P_{i,j} \geq P_{i,j'}, \forall i \in I_j)$ denote that the allocation to (j, j') has a warranted fairness, $\forall j, j' \in J_O$, and $l_{j,j'} = \mathbb{1}(\exists i \in I_j : P_{i,j} < P_{i,j'}) \cdot \mathbb{1}(u_j \leq u_{j'})$ denote that the allocation to (j, j') has a compensatory fairness, $\forall j, j' \in J_O$ according to Definition 1. Here, $\mathbb{1}(\cdot)$ is 1 if its argument is `TRUE` and 0 otherwise.

(i) The *Degree of Fairness* (DoF) is defined as the ratio of the total number of facility pairs (j, j') that are allocated fairly among all the $|J_O| \cdot (|J_O| - 1)$ ordered pairs of distinct open facilities:

$$DoF = \frac{\sum_{j,j' \in J_O, j \neq j'} k_{j,j'} + l_{j,j'}}{|J_O| \cdot (|J_O| - 1)} \in [0, 1].$$

(ii) The *Degree of unwarranted Fairness* (DoF') is defined as the ratio of the total number of facility pairs (j, j') for which an allocation has a compensatory fairness to the total number of facility pairs (j, j') for which an allocation does not have a warranted fairness:

$$DoF' = \frac{\sum_{j,j' \in J_O, j \neq j'} l_{j,j'}}{|J_O| \cdot (|J_O| - 1) - \sum_{j,j' \in J_O, j \neq j'} k_{j,j'}} \in [0, 1].$$

□

Definition 2 states the DoF in terms of ordered pairs of open facilities. Thus, the ordered pair (j, j') might not have a fair allocation even if the pair (j', j) has one. However, Definition 1 implies that at least one of the pairs (j, j') and (j', j) always has a fair allocation, $\forall j, j' \in J_O$. To see this, consider a pair (j, j') that does not have a fair allocation. Then, by Definition 1, $u_j > u_{j'}$ necessarily holds. Thus, $u_{j'} > u_j$ does not hold, and (j', j) is guaranteed to be allocated fairly. Hence, the DoF is at least 0.5 in any feasible solution. The DoF' analogously measures the extent of fairness among facilities that do have competition. Then, Definition 2 defines the DoF' as the odds of the number of facilities with compensatory fair allocations to the odds of those without a warranted fair allocation.

We note that the preceding discussion as well as both Definition 1 and Definition 2 are independent of the underlying optimization models. Different modeling choices—for example, with modifications of the objective function—are expected to give different extents of DoF. Our choice of the objective function in model (1) is governed by this consideration as well. Unfortunately, model (1) still does not guarantee that the optimal allocation has a warranted or compensatory fairness for all pairs of open facilities; i.e., model (1) does not ensure $DoF = 1$ in an optimal solution. See, Example A.5 in Appendix A.3 in this regard. To explain our reasoning, and to further substantiate our choice of the objective function, we now provide one sufficient condition that results in an optimal solution with a perfectly fair allocation. This sufficient condition is a continuous relaxation of the x variables in model (1) as given by the following model.

$$\underline{z}^* = \min_{x,y} \sum_{j \in J} C_j \Big(1 - \frac{\sum_{i \in I} U_i P_{i,j} x_{i,j}}{C_j}\Big)^2 \quad (3a)$$

$$\text{s.t. } (1b) - (1e); (1g); x_{i,j} \geq 0, \forall i \in I, j \in J. \quad (3b)$$

In model (3), we remove the constraints $x_{i,j} \leq 1$ as they are implied by constraints (1e) and (1g). Thus, the only difference in model (1) and model (3) is in constraint (1f): model (1) constrains x as a binary variable, while model (3) provides its continuous relaxation.

THEOREM 1. *An allocation corresponding to an optimal solution to model* (3) *is fair by Definition 1.*

Let (x, y) be an optimal solution to model (3), and $F(x, y)$ denote the objective function of model (3). Using the definition of u from equation (2b), we have $\frac{\partial F}{\partial x_{i,j}} = -2U_i P_{i,j}(1 - u_j), \forall i \in I, j \in J$ and $\frac{\partial F}{\partial y_j} = 0, \forall j \in J$. The necessary Karush-Kuhn-Tucker (KKT) optimality conditions for the non-convex quadratic optimization model (3) are:

- Primal feasibility: constraints (3b)
- Dual feasibility:

$$\alpha_j \frac{U_i P_{i,j}}{C_j} - \nu_i + \mu_{i,j} - \gamma_{i,j} = 2U_i P_{i,j}(1 - u_j) \qquad \forall i \in I, j \in J \tag{4a}$$

$$\beta - \delta_j(1 - 2y_j) = \sum_{i \in I} \mu_{i,j} \qquad \forall j \tag{4b}$$

$$\alpha_j, \beta, \mu_{i,j}, \gamma_{i,j} \geq 0 \qquad \forall i \in I, j \in J. \tag{4c}$$

- Complementary slackness:

$$\alpha_j \left(1 - u_j\right) = 0 \qquad \forall j \in J \tag{5a}$$

$$\beta \left(B - \sum_{j \in J} y_j\right) = 0 \tag{5b}$$

$$\nu_i \left(1 - \sum_{j \in J} x_{i,j}\right) = 0 \qquad \forall i \in I \tag{5c}$$

$$\mu_{i,j} \left(y_j - x_{i,j}\right) = 0 \qquad \forall i \in I, j \in J \tag{5d}$$

$$\gamma_{i,j} x_{i,j} = 0 \qquad \forall i \in I, j \in J \tag{5e}$$

$$\delta_j \left(y_j(1 - y_j)\right) = 0 \qquad \forall j \in J. \tag{5f}$$

We note that these KKT conditions although not sufficient for optimal solutions are necessary; i.e., any optimal solution (x, y) for model (3) must satisfy them. Consider an arbitrary ordered pair (j, j') of distinct open facilities; i.e., $j \neq j'$, $y_j = y_{j'} = 1$. If $P_{i,j} \geq P_{i,j'}$ for every user $i \in I$ assigned to j, then according to Definition 1 the allocation has a warranted fairness. Next, consider that there exists a user $i \in I$ with $x_{i,j} = 1$ with $P_{i,j} < P_{i,j'}$. We show that the allocation has a compensatory fairness.

If $u_{j'} = 1$, then $u_j \leq u_{j'}$ is trivially true, and the allocation to the pair (j, j') has a compensatory fairness. Hence, in the following we assume $u_{j'} < 1$. From equation (1d), $x_{i,j'} = 0$.

From equation (5a), $\alpha_{j'} = 0$, while from equations (5d) and (5e) we further have $\mu_{i,j'} = 0$ and $\gamma_{i,j} = 0$, respectively. Hence, it follows from equation (4a) that

$$-2U_iP_{i,j}(1-u_j) + \alpha_j \frac{U_iP_{i,j}}{C_j} - \nu_i + \mu_{i,j} = 0 \tag{6a}$$

$$-2U_iP_{i,j'}(1-u_{j'}) - \nu_i - \gamma_{i,j'} = 0. \tag{6b}$$

We distinguish two cases below: (i) $u_j = 1$ and (ii) $u_j < 1$.

(i) Assume $u_j = 1$. We prove that this is not possible.

From equation (6) we have

$$\alpha_j \frac{U_iP_{i,j}}{C_j} + \mu_{i,j} = -2U_iP_{i,j'}(1-u_{j'}) - \gamma_{i,j'}.$$

Since $\alpha_j, \mu_{i,j}, \gamma_{i,j'} \geq 0$ and $u_{j'} < 1$, a contradiction follows.

(ii) Next, consider the case that $u_j < 1$. We prove that $u_j < u_{j'}$. We have $\alpha_j = 0$ from equation (5a). Further, since $\mu_{i,j}, \gamma_{i,j'} \geq 0$, equation (6) leads to

$$P_{i,j}(1-u_j) \geq P_{i,j'}(1-u_j').$$

Under the hypothesis, $P_{i,j} < P_{i,j'}$ and $u_j, u_{j'} < 1$. Then, the result follows. □

Theorem 1 shows that a continuous relaxation of the x variables of model (1) alone ensures that all the ordered pairs of facilities have allocations that are fair in an optimal solution of model (3); i.e., there exists an optimal solution with $DoF = 1$ for model (3). In other words, optimal solutions that do not have a fair allocation stem only from the discrete nature of the decision variables x_{ij}. The proof of Theorem 1 hinges on the particular choice of the objective function. Note that model (3) is still non-convex; thus, all solutions that are fair or satisfy the KKT conditions are not always optimal to model (3).

We conclude this section with a disclaimer. The continuous relaxation of binary variables is often a significant modification of any discrete optimization model. In this sense, the above results are of interest purely from a theoretical point of view. That being said, there are three grounds that further warrant our contributions. First, for certain classes of FLPs—in particular the p-median problem—integer solutions are obtained even with a continuous relaxation of both the x and y variables, see, e.g., (ReVelle and Swain 1970, Siegel and Rajaram 2021). Second, the above results are strong in the sense that they provide certificates of fairness despite the non-convexity of the relaxed model and even without requiring the typical additional constraint qualifications. Finally, in Section 5.4 we present computational experiments that provide the DoF and DoF' for model (1) by varying the budget B; these results help determine the shortfall from a DoF of 1.

3. Proportional Fairness

Previous works consider the notion of *proportional fairness* (Huang et al. 2017, Singh 2020); i.e., facilities are assigned users in proportion to their capacities. Huang et al. study a model for allocating discretionary amounts of vaccines to dispensing sites and show that their convex-quadratic optimization model achieves coverage proportional to the weights assigned to sites (Huang et al. 2017). Model (1) significantly differs from the model employed in (Huang et al. 2017) as we also consider the preferences of users and their access to facilities. To this end, we simultaneously maximize access and fairness in an optimal allocation. More importantly, unlike the models considered in the above cited works (Huang et al. 2017, Singh 2020), our models are not convex. The definition of a proportionally fair solution of (Singh 2020) extended to our models is as follows:

Definition 3 (Proportional Fairness). An allocation with utilization defined by equation (2b) is said to be *proportionally fair* if it provides the same utilization for all open facilities. □

Similar to fairness, proportional fairness is also not guaranteed in an optimal solution to model (1); it is easy to construct such examples, see Example A.5 in Appendix A.3. Next, proceeding as in Section 2, we investigate conditions that allow proportional fairness in an optimal solution to model (1). We show that a special case of model (3)—where we entirely eliminate the preferences of user i towards any facility—does indeed ensure proportional fairness in an optimal solution. Thus, as compared to model (1), we require two modifications: (i) we relax the binary restrictions on x to their continuous relaxation as we did in model (3); i.e., $0 \le x_{i,j} \le 1$, and (ii) we remove the preferences of users towards facilities; i.e., $P_{i,j} \leftarrow P_i \in (0,1)$. These modifications are given in the following optimization model.

$$z^*_{PF} = \min_{x,y} \sum_{j\in J} C_j \big(1 - \frac{\sum_{i\in I} U_i P_i x_{i,j}}{C_j}\big)^2 \tag{7a}$$

$$\text{s.t.} \sum_{i\in I} U_i P_i x_{i,j} \le C_j, \forall j \in J; (1c)-(1e); (1g); x_{i,j} \ge 0, \forall i \in I, j \in J. \tag{7b}$$

To prove that an optimal solution of model (7) has a proportionally fair allocation we need the following lemma, the proof of which we reserve for Appendix A.2. Intuitively, the proof rests on the fact that in the absence of preferences towards facilities the total assigned population is $\sum_{j\in J}\sum_{i\in I} U_i P_i x_{i,j} = \sum_{i\in I} U_i P_i$ in every feasible solution, where the equality follows from equation (1d). By allowing a user's population to be split arbitrarily among multiple facilities, the proof shows that given a feasible solution to model (7) we can always construct another feasible solution, where there is a user from which we allocate fractions of population to every utilized facility while maintaining the same utilization of the facilities.

LEMMA 1. *In a feasible solution* (x, y) *to model* (7)*, let* $J_B = \{j \in J : \sum_{i\in I} U_i P_i x_{i,j} > 0\} \subseteq J$*. Consider an* $i^* \in I, j^* \in J$ *such that* $x_{i^*_, j^*} > 0$ *(such a pair exists because of constraint* (1d)*). Then, there exists another feasible solution* (x', y') *that provides the same value of the objective function* (7a) *as* (x, y)*, such that:*

L.(i) $\sum_{i\in I} U_i P_i x'_{i,j} = \sum_{i\in I} U_i P_i x_{i,j}$, $\forall j \in J$

L.(ii) $x'_{i^*_, j} > 0$, $\forall j \in J_B$.

See Appendix A.2. □

COROLLARY 1. *Given an optimal solution* (x, y) *to model* (7) *there exists another optimal solution* (x', y')*, such that conditions L.(i) and L.(ii) of Lemma 1 are satisfied.*

The proof follows directly from Lemma 1. □

THEOREM 2. *An allocation corresponding to an optimal solution to model* (7) *is proportionally fair by Definition 3.*

The proof is similar to that of Theorem 1. Let (x, y) be an optimal solution to model (7). The necessary KKT optimality conditions to model (7) are primal feasibility given by equations (7b), dual feasibility given by equation (4), and complementary slackness given by equation (5), with $P_{ij} \leftarrow P_i \in (0, 1), \forall i \in I, j \in J$. Let $J_O = \{j \in J : y_j = 1\} \subseteq J$ be the set of open facilities. We prove that u_j is the same for all $j \in J_O$. The result is vacuous if $J_O = \emptyset$, while the result holds by definition if $|J_O| = 1$. Hence, consider $|J_O| \geq 2$. We distinguish three cases below.

(i) First, we consider the case that there exists a facility $j \in J_O$ that is fully utilized; i.e., $u_j = 1$. Let $j' \in J_O$ be another facility that is not fully utilized; i.e., $u_{j'} < 1$. We prove this is not possible. Since $u_j > 0$, there exists an i^* such that $x_{i^*_, j} > 0$, hence $\gamma_{i^*_, j} = 0$ from equation (5e). From equation (1d) and $x_{i^*_, j} > 0$ we have $x_{i^*_, j'} < 1$, hence $\mu_{i^*_, j'} = 0$ from equation (5d). Further, $\alpha_{j'} = 0$ from equation (5a). Then, from equation (4a) it follows that $\alpha_j \frac{U_{i^*} P_{i^*}}{C_j} + \mu_{i^*_, j} = \nu_{i^*}$ and $-2U_{i^*} P_{i^*}(1 - u_{j'}) - \gamma_{i^*_, j'} = \nu_{i^*}$. Since $\alpha_j, \mu_{i^*_, j}, \gamma_{i^*_, j'} \geq 0$ and $u_{j'} < 1$ a contradiction follows. Thus, in an optimal solution if there exists even a single facility that is completely utilized, then all other open facilities are also completely utilized.

(ii) Next, we consider the case that there exists a facility $j \in J_O$ that is not utilized at all; i.e., $u_j = 0$. Let $j' \in J_O$ be another facility that is utilized; i.e., $u_{j'} > 0$. We prove this is also not possible. Since $u_{j'} > 0$, there exists an i^* such that $x_{i^*_,j'} > 0$, hence $\gamma_{i^*_,j'} = 0$. Since $u_j = 0$, we have $x_{i^*_,j} = 0$ by the definition of u; further $\alpha_j = 0$ and $\mu_{i^*_,j} = 0$ from equations (5a) and (5d). Then, from equation (4a) it follows that

$$-\nu_{i^*} - \gamma_{i^*_,j} = 2U_{i^*}P_{i^*} \tag{8a}$$

$$\alpha_{j'}\frac{U_{i^*}P_{i^*}}{C_{j'}} - \nu_{i^*} + \mu_{i^*_,j'} = 2U_{i^*}P_{i^*}(1-u_{j'}) \tag{8b}$$

Equation (8) leads to:

$$\alpha_{j'}\frac{U_{i^*}P_{i^*}}{C_{j'}} + \mu_{i^*_,j'} + \gamma_{i^*_,j} = -2U_{i^*}P_{i^*}u_{j'}. \tag{9}$$

Since $\alpha_{j'}, \mu_{i^*_,j'}, \gamma_{i^*_,j} \geq 0$ and $u_{j'} > 0$, a contradiction follows. Thus, in an optimal solution if there exists even a single facility that is open but not utilized at all, then all other open facilities are also not utilized at all.

(iii) Finally, we consider the case that all open facilities have a utilization strictly between 0 and 1. Then, the set $J_B = \{j \in J : \sum_{i\in I} U_i P_i x_{i,j} > 0\} \subseteq J_O$ of utilized facilities is identical to J_O. Thus, from Corollary 1, without loss of generality there exists an $i^* \in I$ such that $x_{i^*_,j} > 0, \forall j \in J_O$; hence, $\gamma_{i^*_,j} = 0, \forall j \in J_O$ from equation (5e), while $y_j = 1, \forall j \in J_O$ from equations (1e) and (1g). Further, since $|J_O| \geq 2$, it follows from equation (1d) that $x_{i^*_,j} < 1, \forall j \in J_O$. Then, $\mu_{i^*_,j} = 0, \forall j \in J_O$ from equation (5d). Lastly, since by hypothesis $u_j < 1, \forall j \in J_O$, we have $\alpha_j = 0, \forall j \in J_O$ from equation (5a).

Hence, from equation (4a) and the definition of u we have

$$\begin{aligned} &\nu_{i^*} = -2U_{i^*}P_{i^*}(1-u_j), && \forall j \in J_O,\\ \text{or,}\quad &u_j = 1 + \frac{\nu_{i^*}}{2U_{i^*}P_{i^*}} < 1, && \forall j \in J_O. \end{aligned}$$

This proves the result. □

Theorem 2 implies that an optimal solution of a modified version of model (1), where preferences for facilities are removed and x is allowed to be fractional, is proportionally fair. This result is again intuitive since without preferences the coverage is independent of the specific facility the population is assigned to. Further, relaxing the binary restrictions on x allows *splitting* the population of a user among several facilities. Finally, we note that cases (i) and (ii) in the proof of Theorem 2 likely do not have a practical relevance. The case of $u_j = 1, \forall j \in J_O$ only holds if $\sum_{i\in I} U_i P_i = \sum_{j\in J_O} C_j$, while the case of $u_j = 0, \forall j \in J_O$ is not even possible when P_i, U_i are strictly positive. In other words, for a realistic model the utilization of all open facilities is equal and strictly between zero and one.

As in Section 2, we conclude this section with another disclaimer. It is tempting to conclude that equal preferences towards facilities renders model (7) impractical. However, such results open several opportunities for exploring characteristics of optimal solutions. The non-convex model (7) that allows fractional assignments—although a significant modification from the original model (1)—still applies to several settings where the population corresponding to user i can be partitioned across facilities via $x_{i,j}$; e.g., splitting according to the last digit of their address or splitting with preferences towards facilities of particular chains. Another theoretical question is the interplay of access and fairness in model (7). We already saw in Theorem 2 that model (7) is perfectly fair towards facilities. Interestingly, in achieving an optimal solution with this proportionally fair allocation, model (7) does not compromise on access. Indeed, the following corollary of Theorem 2 proves that model (7) provides the maximum overall access.

Corollary 2. *An optimal solution to model* (7) *achieves the maximum overall access across any feasible solution.*

The maximum overall access is given by

$$\bar{z}^*_{PF} = \max_{x,y} \sum_{j \in J} \sum_{i \in I} U_i P_i x_{i,j}, \text{s.t. (7b)}.$$

From equation (1d) it follows that in any feasible solution for model (7) we have $\bar{z}^*_{PF} = \sum_{i \in I} U_i P_i$. An optimal solution also achieves this access. □

4. Data Sources and Estimation

In this section, we summarize the data we use for our computational experiments in Section 5. We consider the set of users—$i \in I$—as the set of all the ZIP codes in Bavaria, and the set of facilities—$j \in J$—as the subset of ZIP codes that contain at least one recycling center. Then there are $|I| = 2{,}060$ ZIP codes and 1,529 recycling centers spread over $|J| = 1{,}394$ ZIP codes. As we describe below, we further parameterize the set of facilities as rural or urban. Model (1) requires four parameters: $C_j, U_i, P_{i,j}, \forall i \in I, j \in J$, and B. We solve model (1) for different budgets $B \leq |J|$, and below we briefly describe our methods for estimation of the other three parameters. We reserve the details for Appendix B.

4.1. Estimation of C_j and U_i

We estimate the population of ZIP codes, U_i, using data from suche-postleitzahl.org (2020); see Appendix B.1 for details. We do not have explicit data on the capacities of all the recycling centers, and thus assume that a recycling center's capacity is proportional to the amount of waste that is collected there. To approximate the amount of waste collected, we calculate the "catchment population" of each recycling center; i.e., the number of people that are closest to this recycling center. For details on this estimation, see Algorithm S1 in Appendix B.3. For details on this estimation and the differentiation of rural and urban regions, see Appendix B.2 - B.4.

4.2. Estimation of $P_{i,j}$

Next, we describe our method to estimate the preferences of populations in ZIP code i to recycling centers in ZIP code j. To this end, we construct a data-driven model that includes a total of 618,245 person-trips that determines how far people in Germany travel. Using a least-squares fit on our data, we obtain the fraction of target population willing to travel at least d kilometers that we present in Figure 1. This willingness to travel determines the preferences, $P_{i,j}$. For the model and the details, see Appendix B.5.

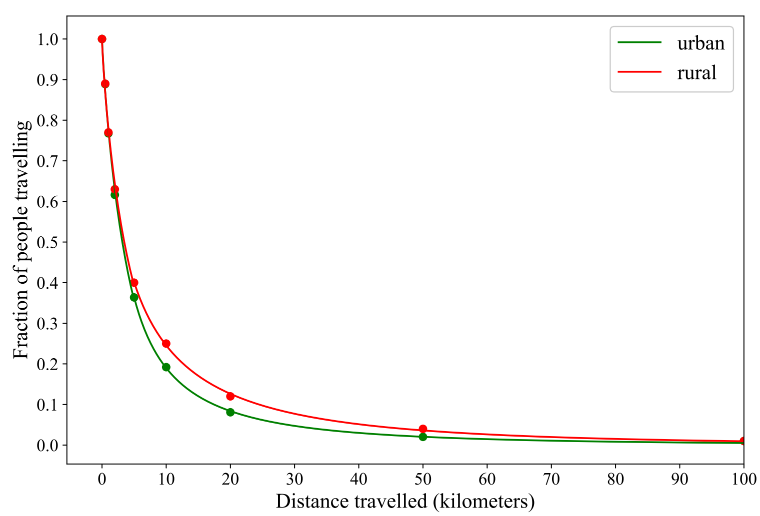


Figure 1 **Willingness to travel curves (solid lines) for urban and rural populations fit to the MiD data (circles) on person-trips for the entire German population. For details, see Section 4.2 and Appendix B.5.**

5. Analysis

5.1. Setup

In this section, we describe the setup for our computational experiments. As we mention in Section 4.1, we consider $|I| = 2{,}060$ ZIP codes and $|J| = 1{,}394$ recycling centers. Thus, there are approximately 2.9 million x variables and as many constraints in equation (1e) alone. To reduce computational effort, we remove (i, j) combinations that have low preferences; this reduction is similar to that implemented in Risanger et al. (2021). Specifically, we set $x_{i,j} = 0, \forall i \in I, j \in J$ for which $P_{i,j} < 0.2$. As a byproduct of this reduction, 27 ZIP codes are left out that do not have a preference of at least 20% to any recycling center. Hence, we implement a post-processing step that assigns these ZIP codes to open recycling centers.

Here, we solve a secondary optimization problem—in addition to model (1)—that considers only the subset $\hat{J} \subset J$ of recycling centers that are open in an optimal solution of the primary model (1). Further, we fix the $x_{i,j}$ variables to the values obtained via the primary model $\forall i \in I \setminus \hat{I}$, where $\hat{I} \subset I$ is the set of ZIP codes that are left unassigned in the primary model. Thus, constraints (1c), (1e), and (1g) are obsolete, and a second computationally easy to solve model results. This secondary optimization model is as follows.

$$\min_{x} \sum_{j \in \hat{J}} C_j \big(1 - \frac{\sum_{i \in I} U_i P_{i,j} x_{i,j}}{C_j}\big)^2 \tag{10a}$$

$$\text{s.t.} \sum_{i \in I} U_i P_{i,j} x_{i,j} \le C_j \qquad \forall j \in \hat{J} \tag{10b}$$

$$\sum_{j \in \hat{J}} x_{i,j} = 1 \qquad \forall i \in \hat{I} \tag{10c}$$

$$x_{i,j} \in \{0, 1\} \qquad \forall i \in \hat{I}, j \in \hat{J}. \tag{10d}$$

Next, to further reduce computational effort, as we mention in Section 2, constraint (1g) can be replaced without loss in optimality by its continuous relaxation. Although there is some data-driven evidence that MIP solvers can generate superior cuts when both x and y are enforced as binary (Ostrowski et al. 2012), in our computational experiments we do not enforce y as binary. For a survey of similar data-driven computational enhancements for MIPs, see, e.g., Klotz and Newman (2013). Finally, we model constraint (1d) as an inequality, $\sum_j x_{i,j} \le 1$, instead of an equality. Then, we use model (10) as a post-processing step to assign the previously unassigned ZIP codes. We denote these two versions as "Strict" (with constraint (1d)) and "Loose" (with constraint (1d) implemented as $\sum_{j \in J} x_{i,j} \le 1$). Both versions include the secondary optimization model (10).

The above two enhancements lead to solutions that are feasible to model (1) but suboptimal. We quantify the tradeoff between suboptimality and the computational savings in

Section 5.3. With these two enhancements, we carry out all computational experiments on two high performance computing clusters with Intel Xeon E5-2643 v4 processors with 256 GB of RAM, Pyomo version 6.1.2 and Gurobi version 9.1.2. We seek to solve all optimization models to an optimality tolerance of 0.5%, except model (10) to a 0% tolerance.

5.2. Analysis

Budget	Regions	Overall access [%]	Travel distance [km]			Utilization [%]		
			p10	p50	p90	p10	p50	p90
100%	all	62.2	0.7	2.0	6.0	27.0	40.5	56.7
	rural	61.6	0.7	2.1	6.6	26.1	39.8	55.4
	urban	62.8	0.6	1.9	4.8	28.7	43.0	59.2
90%	all	61.9	0.7	2.2	6.4	29.2	42.3	58.0
	rural	61.1	0.7	2.4	7.0	28.6	41.7	56.8
	urban	62.7	0.6	2.0	4.8	31.2	43.9	60.5
80%	all	61.2	0.7	2.5	7.0	31.3	43.9	59.7
	rural	60.1	0.7	2.8	7.6	30.9	43.6	58.8
	urban	62.4	0.7	2.2	5.0	32.8	44.5	62.0
70%	all	60.4	0.8	2.9	7.5	32.7	46.4	62.3
	rural	58.8	0.8	3.4	8.1	32.4	46.5	61.8
	urban	62.1	0.8	2.2	5.5	33.4	46.0	62.8
60%	all	59.1	0.9	3.4	8.0	33.7	47.7	63.7
	rural	57.0	1.0	4.1	8.8	33.0	48.2	63.2
	urban	61.4	0.9	2.4	5.8	34.6	46.9	64.6
50%	all	57.6	1.0	4.1	8.9	34.4	49.2	65.8
	rural	54.5	1.2	5.1	9.9	34.2	50.2	65.9
	urban	60.8	0.9	2.6	6.7	34.6	47.2	65.8
40%	all	55.4	1.2	5.0	9.8	36.0	51.3	69.4
	rural	51.6	1.4	6.0	10.8	35.2	53.3	69.6
	urban	59.2	1.0	3.0	7.8	36.2	47.5	66.6
30%	all	52.6	1.5	6.0	11.5	37.7	55.4	73.0
	rural	48.2	1.8	6.8	11.9	37.5	58.2	74.6
	urban	57.1	1.1	3.8	8.5	38.0	52.3	71.4

Table 1 Estimated overall access, travel distances, and utilization of open facilities for different budgets for rural, urban, and all regions of Bavaria on solving model (1). Here, the "Budget" column denotes the percentage of open facilities. The "p10", "p50", and "p90" columns denote the 10^{th}, median, and 90^{th} percentiles, respectively. For details, see Section 5.2.

First, we present our results on solving model (1) for different budgets, B; i.e., the number of recycling centers that are open. Fig. 2a shows the overall access, given by $100\,\frac{\sum_{j\in J} a_j}{\sum_{i\in I} U_i}$, for different budgets of open recycling centers, while Fig. 2b shows the travel distances from the ZIP codes to the assigned recycling centers in an optimal solution. For increasing budget, the overall access increases while the median travel distance decreases. Table 1 presents these results. The maximum overall access is no more than 62.2%, even when all the recycling centers are open. This value is determined by the stringent preferences, as we discussed in

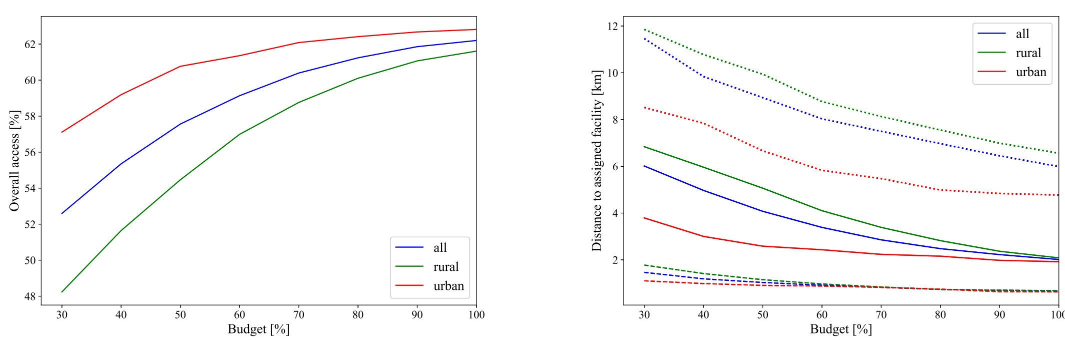
Section 2, given by the travel model in equation (16), and the relatively low value of the maximum access is primarily due to two reasons.

(i) First, several regions in Bavaria have a sparse number of recycling centers. For example, there are only two recycling centers in the district of Schweinfurt with about 116,000 inhabitants spread over an area of 841,000 square kilometers. Similarly, the district of Tirschenreuth with a population of 72,000 and an area of 1,084,000 square kilometers has only one recycling center. Thus, even when all the recycling centers are open, there is poor access, and the travel distances are large.

(ii) Second, our travel model (16) is a conservative estimate of the preferences to recycling centers. This is because our model is based on data that shows how far people actually travel for errands instead of how far they would be *willing* to travel. Similar underestimates in travel models have been reported before; see, e.g., Risanger et al. (2021). In other words, the actual access to the recycling centers might be larger than that predicted by our results; however, if our estimation is conservative consistently across our data, the choice of the optimal recycling centers would not change.

In our computational experiments, we distinguish our results for rural and urban regions after optimizing for all the regions. For the overall access and travel distances this distinction is with respect to the users' location, while for the utilization of open facilities it is with respect to the facilities' location. As demonstrated in Fig. 1, users in rural regions have slightly larger preferences to the same facility than those in urban regions. However, as we observe from Table 1, access is always lower in rural regions than in urban regions for all budgets; correspondingly, the travel distances in the rural regions are larger. The reasons for this are subtle and are as follows. Although, for the same distance, the preferences are larger for users in rural regions than those in urban regions, the distances to the closest recycling centers are also, on average, larger in the rural regions. For users in rural regions the average preference to the closest open recycling center when all the facilities are open is slightly less (0.61) than that for users in urban regions (0.63). Equivalently, the average distance to the closest open recycling center is larger for users in rural regions (3.0 kms) than that in urban regions (2.3 kms). Consequently, the access to the assigned recycling center is lower for users in rural regions. Secondly, for small budgets, model (1) favors larger facilities to open. The average capacity of a facility in a rural region is 10,905 persons while that of one in an urban region is 22,865 persons. For $B = 558$ (or 40%), the largest 139 facilities are all opened. The share of large facilities among facilities in urban regions is higher than among facilities in rural regions; e.g., for a 40% budget ($B = 40\%|J|$), in an optimal solution 55% of *all* the facilities that are located in urban regions are open, while only 34% of *all* the facilities that are located in rural regions are open. This is despite the fact that there are significantly more facilities in rural regions than in urban regions; 70% of all the 1,394 facilities are in rural regions. For a 40% budget, in an optimal solution, only 58% (42%) of the opened facilities are in rural (urban) regions. Thus, users in rural regions are forced to travel farther to facilities outside of their ZIP codes. Indeed, for a 40% budget, 58% of the users in rural regions are assigned to a facility that is not their closest open facility. In contrast only 38% of the users in urban regions are assigned to facilities that are not the closest.

The above discussion demonstrates that the burden of providing access to rural regions falls predominantly on the fewer rural facilities. This results in larger utilization of the rural facilities and a lower access. However, the gaps between the rural and urban regions shrink as more facilities are allowed to be open. The overall access for a 40% budget deviates by 7.6 percentage points (i.e., the urban populations have 7.6 percentage points more access than the rural populations) while for a 90% budget the difference drops to only 1.6 percentage points. Similarly the deviation in the median utilization changes from -5.8 percentage points to 2.2 percentage points (i.e., the urban facilities are utilized 2.2 percentage points more than the rural facilities) for an increase in budget from 40% to 90%. For details, see Table 1.

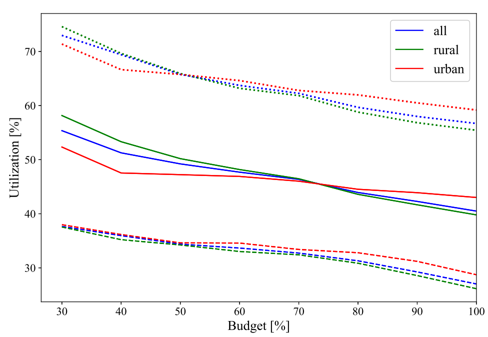


(a) Overall access

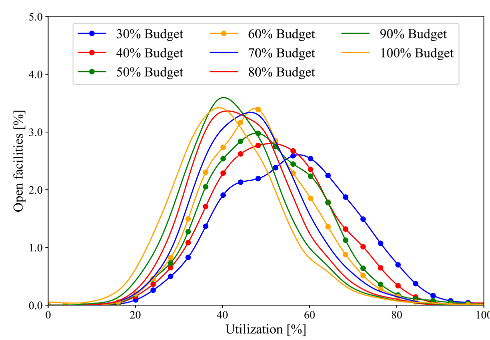
(b) Travel distance

Figure 2 **Estimated overall access (Fig. 2a) and the corresponding travel distances (Fig. 2b) for different budgets for rural (in red), urban (in green), and all (in blue) regions of Bavaria on solving model (1). "Budget" denotes the percentage of open facilities. In Fig. 2b, the solid, dashed, and dotted lines denote the median, 10th percentile and 90th percentile of the travel distances, respectively. For details, see Section 5.2.**

The median utilization of the open facilities drops with increasing budget, since the burden of providing access is distributed among more recycling centers, see Fig. 3a. When all the recycling centers are open—which provides the maximum access of 62.2%—the median utilization is at its lowest value of 40.5%, with a wide range of the 10^{th} and 90^{th} quantiles, see Table 1. However, although the marginal gains in access by opening more facilities steadily drop, the same is not true for the marginal drops in the median utilization. The increase in overall access by opening the last 30% facilities (i.e., varying the budget between 70% and 100%) is less than 2 percentage points, while the decrease in median utilization is almost 6 percentage points. These observations also hold true when considering rural or urban facilities alone.



(a) Utilization of open facilities



(b) Distribution of the utilization of open facilities

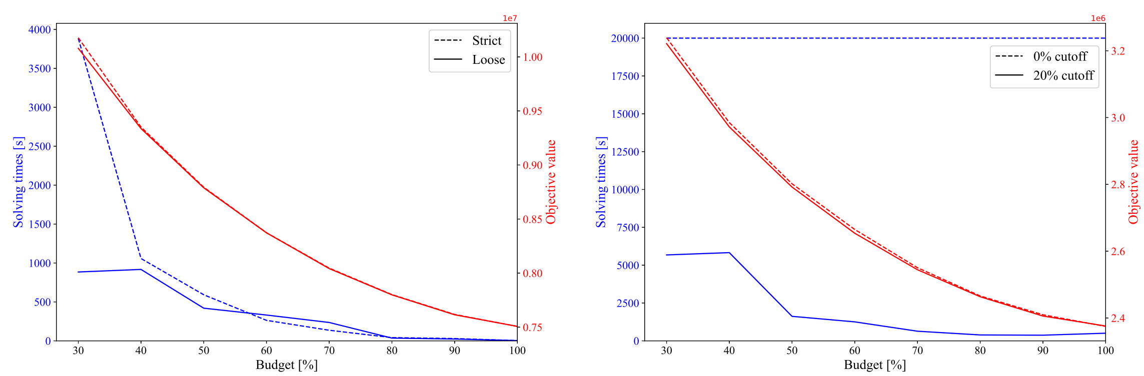
Figure 3 Optimal utilization of the open facilities for different budgets. "Budget" denotes the percentage of open facilities. Fig. 3a provides the utilization for the rural (in red), urban (in green), and all (in blue) regions of Bavaria on solving model (1). The solid, dashed, and dotted lines denote the median, 10th percentile and 90th percentile of the utilization, respectively. Fig. 3b shows the distribution of utilization for budgets, where the y-axis provides the relative frequency of the open facilities. A broad distribution suggests a large dispersion. For analogous figures on the distribution of the utilization of facilities in rural and urban areas, see Fig. S2a and Fig. S2b in Appendix C, respectively. For details, see Section 5.2.

5.3. Effect of computational enhancements

In this section, we provide a brief analysis of the tradeoff between the suboptimality of the solution with the two computational enhancements we mention in Section 5.1. First, we compare the quality of an optimal solution provided by the Loose and Strict models. Both of these versions involve a post-processing step due to the 20% cutoff. A few ZIP codes contain only a very small number of facilities to which the respective travel probability exceeds 20%;

this can result in infeasibility of the Strict version. The Loose version is always feasible as it allows some ZIP codes to be left unassigned; there are 76 such ZIP codes (i.e., 3.69%) for a 30% budget, while there are no such unassigned ZIP codes for a 100% budget. In addition, in both the versions, 27 ZIP codes (i.e., 1.31%) are excluded as they fall out of the 20% cut-off. After the post-processing step all ZIP codes are assigned.

We observe that for lower budgets, more ZIP codes remain unassigned via the primary model. In other words, the Loose version becomes increasingly accurate for larger budgets. Table 2 summarizes these observations for a 30% budget. Further, Fig. 4a demonstrates that the objective function values of the two models are close, however we save significant computational effort in the Loose model. The largest deviation between the two objective function values is 0.96% for a budget of 30%; for the same budget, the improvement in the runtime is 77.37%. On average, the objective function values obtained by the Loose model differ only by 0.16%, suggesting that the significant reduction in computational effort comes at the expense of at most a marginal deterioration in the quality of solutions.



(a) Effect of loose inequality (b) Effect of cutoff

Figure 4 **Tradeoff between the computational runtimes (blue) and objective function values (red) for model (1). In Fig. 4a, the "Strict" and "Loose" denote the two versions when constraint (1d) is implemented as an equality and inequality, respectively. Both versions include a 20% cutoff and a post-processing step. In Fig. 4b, the "0% cutoff" and "20% cutoff" denote the two versions without and with a 20% cutoff, respectively. Both versions include constraint (1d) as an inequality. The total number of users and facilities in Fig. 4b are 30% of that in Fig. 4a. For details, see Section 5.2.**

Next, we compare the quality of an optimal solution when we remove all (i, j) combinations that have preferences less than 20%. Instances with the full set of users and facilities could not be generated without this cutoff (see Fischetti et al. (2017) for one standard way to resolve this); hence, we only consider instances where the first 30% of users and facilities are selected. Then, we solve model (1) twice — first without removing any (i, j) combinations, and second, with removing all combinations with $P_{ij} < 20\%$; we denote these versions as "0% cutoff" and "20% cutoff", respectively. Fig. 4b presents our results. The 0% cutoff implementation could not be solved to optimality within a time limit of 20,000 seconds for all budgets we consider. On the other hand, the 20% cutoff implementation is solved to optimality with significantly lesser effort. Again, the deviation from optimality is marginal. The conservative gap between the best feasible solution reported by the 20% cutoff and the best bound reported by the 0% cutoff is, on average, 2.4%. For a detailed analysis with a variety of instances, see Table S1 in Appendix C.

	Strict		**Loose**		**Improvement**
	Before	After	Before	After	
Assigned zip codes [%]	98.69	100.00	95.00	100.00	0.00%
Overall access [%]	52.34	52.43	52.30	52.60	-0.31%
Objective value	10,195,018.2	10,177,542.3	10,124,222.9	10,079,376.8	0.96%
Solving time [s]	3,872.8	3,873.2	871.0	876.4	77.37%

Table 2 **Comparison of the quality of solutions for model (1) when constraint (1d) is implemented as an equality ("Strict") and as an inequality ("Loose") for a 30% budget. Both versions include a 20% cutoff and a post-processing step. The "Improvement" column shows the relative difference between the two "After" columns; i.e., $100\frac{\text{Strict-Loose}}{\text{Strict}}$. The solving time for the Loose model is 77.37% better at a deviation of 0.31% in overall access and 0.96% in the objective function value. For details, see Section 5.3.**

5.4. Tradeoff between accessibility and fairness

As we mention before, the objective function (1a) of model (1) seeks to simultaneously achieve the conflicting goals of maximizing overall access and fairness in utilization of the open recycling centers. Theorem 2 shows that a modification of model (1)—model (7)—guarantees perfect proportional fairness. In Example A.5, we demonstrate that proportional fairness does not necessarily hold for model (1). In this section, we examine the extent to which proportional fairness is achieved via model (1) by studying various measures of the variability in the utilization of the open recycling centers.

First, Table 1 provides the median, 10^{th}, and 90^{th} percentiles of the utilization, while Fig. 3b illustrates the distribution of the utilization of open facilities for various budgets. The distributions are consistently unimodal; for larger budgets they exhibit a positive skewness characterized by a long right tail. Although the gaps between the 10^{th} and 90^{th} percentiles are significant, the distributions become narrower for larger budgets suggesting the dispersion decreases with increasing budget. Further, the peaks of the distributions shift to the left reiterating the notion that the utilization decreases with increasing budget, as we discuss in Section 5.2.

An optimal solution of model (1) that achieves perfect proportional fairness, as characterized by Theorem 2, has zero variance of the utilization of open facilities. Since the objective function (1a) of model (1) seeks to achieve low variability by minimizing the *capacity-weighted* sum of squared idleness $1 - u_j$, the *weighted* variance of the utilization of open facilities is one measure to determine the extent to which proportional fairness is achieved. Here, the weights correspond to the capacities of the facilities. We standardize this measure by dividing the weighted standard deviation by the weighted average utilization of open facilities, and obtain the *weighted coefficient of variation* (CV_w) as $CV_w = \frac{\sqrt{Var_w(u)}}{\bar{u}_w}$. Here, $\bar{u}_w = \frac{\sum_{j \in J_O} C_j u_j}{\sum_{j \in J_O} C_j}$ and $Var_w(u) = \frac{\sum_{j \in J_O} C_j (u_j - \bar{u}_w)^2}{\sum_{j \in J_O} C_j}$ are the weighted average and weighted variance of the utilization of open facilities, respectively, and $J_O = \{j \in J : y_j = 1\}$ is the set of open facilities in a feasible solution. We distinguish these results according to the region in which the facilities are located. For all budgets in Table 3, the CV_w is at most just above 30% for both the rural and urban regions.

The third measure we examine is the DoF and the DoF' as given by Definition 2. The last two columns in Table 3 present these values. The DoF is nearly one for all budgets we consider. This uniform behavior is because the allocation has a warranted fairness for nearly all of the open facility pairs. In other words, model (1) drives towards solutions that lower any basis for arguments between facilities. Further, the DoF' column demonstrates that among facility pairs for which the allocation has a warranted fairness, over two-thirds do have allocations that are compensatory fair. With increasing budget, the DoF increases; i.e., the fraction of pairs with an unfair allocation shrinks. Similarly, with increasing budget, the DoF' generally decreases; i.e., the fraction of pairs to which the allocation has a compensatory fairness shrinks even faster than the fraction of pairs with an unfair allocation.

The fraction of open recycling centers whose utilization lies within one (two/three) standard deviation(s) of the mean is above 64% (95%/99.3%) for all budgets. This behavior is similar to that observed when the data is normally distributed, where roughly 68%, 95% and 99.7% of values lie within one, two and three standard deviation(s) of the mean, respectively,

Budget	**Regions**	$\bar{u}_w$ [%]	$\sqrt{Var_w(u)}$ [%]	CV_w [%]	DoF [%]	DoF' [%]
100%	all	40.23	11.23	27.91	100.00	67.54
	rural	38.74	11.81	30.48		
	urban	41.86	10.31	24.64		
90%	all	40.93	10.86	26.53	100.00	76.38
	rural	39.72	11.45	28.82		
	urban	42.22	10.03	23.75		
80%	all	41.74	10.78	25.83	100.00	83.92
	rural	40.85	11.55	28.26		
	urban	42.67	9.85	23.08		
70%	all	42.62	10.93	25.65	100.00	88.80
	rural	42.19	11.80	27.96		
	urban	43.04	9.97	23.17		
60%	all	43.57	11.10	25.47	99.99	89.41
	rural	43.43	12.27	28.26		
	urban	43.70	9.84	22.52		
50%	all	44.71	11.57	25.87	99.99	89.40
	rural	45.34	12.97	28.61		
	urban	44.15	10.13	22.93		
40%	all	46.09	11.93	25.88	99.99	92.92
	rural	47.57	13.49	28.37		
	urban	44.85	10.28	22.92		
30%	all	48.59	12.93	26.61	99.99	94.32
	rural	51.08	14.72	28.82		
	urban	46.63	10.94	23.45		

Table 3 **The third to fifth columns present the weighted mean ($\bar{u}_w$), weighted standard deviation ($\sqrt{Var_w(u)}$), and weighted Coefficient of Variation (CV_w) of the utilization of open recycling centers for different budgets, respectively. The last two columns present the Degree of Fairness (DoF) and the Degree of unwarranted Fairness (DoF') as defined in Definition 2. For details, see Section 5.4.**

see e.g. (Casella 2002, Chapter 3). Finally, we find a strong linear correlation between the capacity of an open recycling center j, C_j, and its accessibility, a_j, given by equation (2a). Using linear regression, we obtain linear functions $a_j = m \cdot C_j$ that fit the data well; for all budgets we consider, the respective R^2-values of these fits are greater than 0.9. The fits are better for large budgets; for a 70% budget or larger, the R^2-values are greater than 0.94. For a visualization of the linear correlation between a and C for an exemplary budget of 30%, see Fig. S3 in Appendix B.

The above observations provide data-driven evidence that model (1) achieves proportional fairness to a level that is not compromised by maximizing access. However, achieving this level of fairness does require compromising some degree of access. To determine the extent of the sacrifice paid in terms of access, we compare the overall access achieved by model (1) with the *maximum* overall access that is possible. Consider the following optimization model.

$$z^*_{MA} = \max_{x,y} \sum_{j\in J}\sum_{i\in I} P_{i,j}U_i x_{i,j}, \text{s.t. (1b)} - \text{(1g)}. \tag{11a}$$

Model (11) does not consider fairness, and only seeks to maximize overall access using the same set of constraints as model (1). Bertsimas et al. introduce the notion of the *price of fairness*, and define it as the relative reduction of the sum of utilities achieved by a fair solution compared to a solution that maximizes these utilities (Bertsimas et al. 2011). We adapt this notion for our work by computing the price of fairness as the relative reduction of the optimal access achieved by model (1) compared with the maximum possible access achieved by model (11) as follows:

$$\text{Price of fairness} = \frac{\text{maximum overall access} - \text{optimal overall access}}{\text{maximum overall access}}. \tag{12}$$

In equation (12), the maximum overall access is given by z^*_{MA} from model (11), while the optimal overall access is given by $\sum_{j\in J} a_j$ from equation (2a) and model (1). Table 4 summarizes the results for all budgets we consider; the price of fairness is at most 2%. The above analysis provides additional data-driven evidence that model (1) is well-equipped in providing good accessibility while simultaneously ensuring a large extent of fairness.

Budget	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
Optimal overall access [%]	52.6	55.4	57.6	59.1	60.4	61.2	61.9	62.2
Maximum overall access [%]	53.7	56.4	58.3	59.8	60.9	61.5	62.2	62.4
Price of fairness [%]	2.0	1.8	1.3	1.1	0.9	0.4	0.6	0.3

Table 4 Comparison of optimal and maximum overall access for different budgets. The overall access is defined by $\sum_{j\in J} a_j$, where a_j is given by equation (2a). The optimal overall access is that achieved by model (1), while the maximum overall access is that achieved by model (11). The price of fairness is the relative reduction of the optimal overall access compared with the maximum overall access given by equation (12). For details, see Sections 5.4.

5.5. Summary

Sections 4 - 5.4 address the applied component of our theoretical work in Sections 2 - 3. Here, we provide some policy implications of our results for Bavaria that have standalone value in their own regard. Recent years have seen a decline in the number of operational recycling centers in Bavaria (Bayerisches Landesamt für Umwelt 2020). Data-driven models, such as those we present, can help provide informed choices for such closures. Our analysis provides insight into the extent to which waste management services operate, even when a significant number of recycling centers are closed. Policymakers have also recognized that the success of recycling centers is driven by the degree of usage by the public (Bayerisches Landesamt für Umwelt 2015). Smartly locating operational recycling centers is one tool to streamline usage by the public. Efficient public informational campaigns is another, and policymakers are implementing steps in this regard (Bayerisches Landesamt für Umwelt 2015).

Our analysis demonstrates that although the marginal loss in overall access by closing recycling centers is quite small, fairness is severely impacted for facilities. Further, the marginal drop in overall accessibility to open facilities is significantly larger among facilities in rural regions than in urban regions. Further increasing this disparity, the additional burden imposed by closing recycling centers is not shared uniformly by facilities—the responsibility to ensure access falls slightly more on facilities in rural regions. Rural regions have more facilities with smaller capacities. In a choice between a closure of a larger and smaller facility, a policymaker might prefer closing the smaller facility. Indeed, our models provide evidence to this — to achieve high degrees of access, closing a smaller facility is less damaging. Consequently, facilities that do remain open in the rural regions must shoulder a larger burden of users. Such ethically challenging policy decisions that lead to increased disparities between regions can be backed up by policymakers with quantitative evidence, such as those provided by our models.

Our analysis relies on several simplifying assumptions. We parameterize recycling centers only by their capacity and their location. However, not all facilities accept the full spectrum of recyclable waste. Further, recycling centers in Germany impose different fees as well as limits on the accepted quantities of each type of waste, see, e.g., (Bayerisches Staatsministerium für Umwelt und Verbraucherschutz and Bayerisches Landesamt für Umwelt 2021). These distinctions offer further opportunities to provide a higher-fidelity classification of recycling centers. Additionally, the deposited waste is not always processed on the site, but is often transported further for a final disposition. Our work ignores the transportation costs associated with such transfers; we also ignore any terminal costs for closures. Further, the travel distances used as input to determine the preferences are simplified estimates. First, we assume the entire population of a ZIP code resides at its centroid. Second, we use geodesic distances — the shortest distances on the surface of the earth between the ZIP codes and the recycling centers — rather than the longer road distances. These two simplifications are frequently used when measuring the distances of two geographic objects, however they can lead to underestimates in the actual distance traveled by residents (Jones et al. 2010).

6. Conclusions

We propose an optimization framework for the problem of assigning users to undesirable facilities, while balancing accessibility of the users and fairness of the usage of facilities. Our central model is a non-convex quadratic-binary program with constraints similar to a p-median problem, but a significantly different objective function. With sequential relaxations, we present analytical results, based on the KKT optimality conditions of the underlying optimization models, that help determine the tradeoff between access and fairness. We demonstrate how our work extends several concepts of fairness that already exist in the literature. We further present metrics that measure the extent to which fairness is achieved in feasible solutions. Then, we present an application of our work to allocate residents to recycling centers, using data from the state of Bavaria.

Our work offers several grounds for extensions in the future. First, uncertainty could be accounted for in the preferences of users towards facilities. Then, a probability distribution governs the preferences and stochastic optimization models could be employed. Second, analytical or greedy lower and upper bounds for model (1) can be developed. Future work could also focus on a heuristic that incorporates criteria for both fairness and accessibility to further improve the upper bounds.

All our data, models, codes, and an appendix are publicly available at `https://github.com/schmitt-hub/preferential_access_and_fairness_in_waste_management`.

Acknowledgments

We thank David Morton for his suggestions and comments. We gratefully acknowledge the compute resources and support provided by the Erlangen Regional Computing Center (RRZE). Bismark Singh was partially supported by the Bavarian State Ministry for Science and Art (Bayerisches Staatsministerium für Wissenschaft und Kunst) under the project "Greedy algorithms for fair allocations and efficient assignments within facility location optimization problems".

References

Bayerisches Landesamt für Umwelt, ed. (2015) *Wertstoffhof 2020 - Getrennthaltungsgebot und Novelle des ElektroG*, UmweltSpezial, URL `https://www.bestellen.bayern.de/application/eshop_app000009?SID=62794461`.

Bayerisches Landesamt für Umwelt (2020) Hausmüll in Bayern - Bilanzen 2019: Informationen aus der Abfallwirtschaft. URL https://www.abfallbilanz.bayern.de/doc/2019/Abfallbilanz2019.pdf.

Bayerisches Staatsministerium für Umwelt und Verbraucherschutz, Bayerisches Landesamt für Umwelt (2021) Abfallratgeber Bayern. URL https://www.abfallratgeber.bayern.de/.

Berman O, Drezner Z, Tamir A, Wesolowsky GO (2009) Optimal location with equitable loads. *Annals of Operations Research* 167(1):307–325, ISSN 0254-5330, URL http://dx.doi.org/10.1007/s10479-008-0339-9.

Bertsimas D, Farias VF, Trichakis N (2011) The price of fairness. *Operations Research* 59(1):17–31, ISSN 0030-364X, URL http://dx.doi.org/10.1287/opre.1100.0865.

Cappanera P (1999) A survey on obnoxious facility location problems. Technical report, University of Pisa, Pisa, URL http://eprints.adm.unipi.it/2014, Technical Report: TR-99-11.

Casella G (2002) *Statistical inference* (Australia Pacific Grove, CA: Thomson Learning), 2 edition, ISBN 0534243126.

Erkut E, Neuman S (1989) Analytical models for locating undesirable facilities. *European Journal of Operational Research* 40(3):275–291, ISSN 0377-2217, URL http://dx.doi.org/10.1016/0377-2217(89)90420-7.

Fischetti M, Ljubić I, Sinnl M (2017) Redesigning benders decomposition for large-scale facility location. *Management Science* 63(7):2146–2162, URL http://dx.doi.org/10.1287/mnsc.2016.2461.

Huang HC, Singh B, Morton DP, Johnson GP, Clements B, Meyers LA (2017) Equalizing access to pandemic influenza vaccines through optimal allocation to public health distribution points. *PLOS ONE* 12(8):e0182720, URL http://dx.doi.org/10.1371/journal.pone.0182720.

Ibrahim MA (2020) Risk of spontaneous and anthropogenic fires in waste management chain and hazards of secondary fires. *Resources, Conservation and Recycling* 159:104852, ISSN 0921-3449, URL http://dx.doi.org/10.1016/j.resconrec.2020.104852.

Jones SG, Ashby AJ, Momin SR, Naidoo A (2010) Spatial implications associated with using Euclidean distance measurements and geographic centroid imputation in health care research. *Health Services Research* 45(1):316–327, URL http://dx.doi.org/10.1111/j.1475-6773.2009.01044.x.

Kaplan S (1974) Application of programs with maximin objective functions to problems of optimal resource allocation. *Operations Research* 22(4):802–807, URL http://dx.doi.org/10.1287/opre.22.4.802.

Kaza S, Yao LC, Bhada-Tata P, van Woerden F (2018) *What a waste 2.0: A global snapshot of solid waste management to 2050* (Washington, DC: World Bank), ISBN 978-1-4648-1329-0, URL http://dx.doi.org/10.1596/978-1-4648-1329-0.

Kelly FP, Maulloo AK, Tan DKH (1998) Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society* 49(3):237–252, ISSN 0160-5682, URL http://dx.doi.org/10.1057/palgrave.jors.2600523.

Klotz E, Newman AM (2013) Practical guidelines for solving difficult linear programs. *Surveys in Operations Research and Management Science* 18(1-2):1–17, URL http://dx.doi.org/10.1016/j.sorms.2012.11.001.

Marín A (2011) The discrete facility location problem with balanced allocation of customers. *European Journal of Operational Research* 210(1):27–38, ISSN 0377-2217, URL http://dx.doi.org/10.1016/j.ejor.2010.10.012.

Marsh MT, Schilling DA (1994) Equity measurement in facility location analysis: A review and framework. *European Journal of Operational Research* 74(1):1–17, ISSN 0377-2217, URL http://dx.doi.org/10.1016/0377-2217(94)90200-3.

Morell D (1984) Siting and the politics of equity. *Hazardous Waste* 1(4):555–571, ISSN 0738-6168, URL http://dx.doi.org/10.1089/hzw.1984.1.555.

Ostrowski J, Anjos MF, Vannelli A (2012) Tight mixed integer linear programming formulations for the unit commitment problem. *IEEE Transactions on Power Systems* 27(1):39–46, URL http://dx.doi.org/10.1109/tpwrs.2011.2162008.

Pirkul H, Schilling DA (1988) The siting of emergency service facilities with workload capacities and backup service. *Management Science* 34(7):896–908, URL http://dx.doi.org/10.1287/mnsc.34.7.896.

Plastria F, Carrizosa E (1999) Undesirable facility location with minimal covering objectives. *European Journal of Operational Research* 119(1):158–180, ISSN 0377-2217, URL http://dx.doi.org/10.1016/S0377-2217(98)00335-X.

ReVelle CS, Swain RW (1970) Central facilities location. *Geographical Analysis* 2(1):30–42, URL http://dx.doi.org/https://doi.org/10.1111/j.1538-4632.1970.tb00142.x.

Risanger S, Singh B, Morton D, Meyers LA (2021) Selecting pharmacies for COVID-19 testing to ensure access. *Health Care Management Science* 24(2):330–338, ISSN 1386-9620, URL http://dx.doi.org/10.1007/s10729-020-09538-w.

Siegel Z, Rajaram K (2021) p-median problems and solution strategies. techreport, University of California, Los Angeles, URL https://zsiegel92.github.io/writing_repo/UCLA/mgmt242/pmedian.pdf.

Singh B (2020) Fairness criteria for allocating scarce resources. *Optimization Letters* 14(6):1533–1541, URL http://dx.doi.org/10.1007/s11590-020-01568-1.

suche-postleitzahlorg (2020) Karten von Deutschland. URL https://www.suche-postleitzahl.org/plz-karte-erstellen.

Tversky A, Simonson I (1993) Context-dependent preferences. *Management Science* 39(10):1179–1189, URL http://dx.doi.org/10.1287/mnsc.39.10.1179.

Wolsey LA (1998) *Integer programming.* A Wiley-Interscience Publication (New York, NY: Wiley), ISBN 978-0-471-28366-9.

Yekeen TA, Xu X, Zhang Y, Wu Y, Kim S, Reponen T, Dietrich KN, Ho SM, Chen A, Huo X (2016) Assessment of health risk of trace metal pollution in surface soil and road dust from e-waste recycling area in China. *Environmental Science and Pollution Research International* 23(17):17511–17524, URL http://dx.doi.org/10.1007/s11356-016-6896-6.