Individual and group fairness in geographical partitioning

Ilya O. Ryzhov   John Gunnar Carlsson   Yinchu Zhu

November 26, 2025

Abstract

Socioeconomic segregation often arises in school districting and other contexts, causing some groups to be over- or under-represented within a particular district. This phenomenon is closely linked with disparities in opportunities and outcomes. We formulate a new class of geographical partitioning problems in which the population is heterogeneous, and it is necessary to ensure fair representation for each group at each facility. We prove that the optimal solution is a novel generalization of the additively weighted Voronoi diagram, and we propose a simple and efficient algorithm to compute it, thus resolving an open question dating back to Dvoretzky et al. (1951). The efficacy and potential for practical insight of the approach are demonstrated in a realistic case study involving seven demographic groups and 78 district offices.

1 Introduction

Consider a service system in which individuals are served by facilities at different locations within a geographical region. For example, the facilities could represent schools, polling places, or commercial fulfillment centers. The geographical partitioning problem (Carlsson & Devulapalli 2013) divides the region into non-overlapping districts, such that all individuals residing in the same district are served by the same facility. The goal is to choose a partition that optimizes some measure of social welfare, most commonly the average travel cost per individual (Carlsson et al. 2016).

We formulate and study a novel variant of this problem where the population is heterogeneous, consisting of multiple demographic groups, each with a different spatial distribution throughout the region. Again we optimize the expected cost, but now we also impose a new *group fairness* condition: each subpopulation can be neither over- nor under-represented at any facility. In other words, the districts are designed in such a way that the proportion of the population belonging to a particular group in any district must match that group's incidence in the entire population. This condition is also known as "demographic parity" in the literature (Dwork et al. 2012).

We are motivated by the well-known phenomenon of socioeconomic segregation in school dis-

tricts (Reardon & Bischoff 2011). As income inequality increased over the past several decades, both high- and low-income households exhibited a greater tendency to cluster together geographically, resulting in segregation not only between school districts but also between individual schools in the same district (Owens et al. 2016). Much of this behavior has been driven by families with children (Owens 2016), for whom school boundaries are a major factor in residential decisions. Income segregation between school districts has been linked to gaps in student achievement (Ryabov & Van Hook 2007, Owens 2018), as many high-income districts receive more resources due to higher revenues from local taxes (Baker et al. 2017). Unequal representation of different socioeconomic groups across districts is thus closely linked with disparities in opportunities and outcomes.

The causes of income segregation are complex and, generally, beyond the scope of what can be "fixed" by a single optimization model. Nonetheless, incorporating demographic parity into districting decisions could be seen as one possible way to remove a perceived incentive for economic segregation. In this connection, it is interesting that Saporito (2017a,b) found empirical evidence that segregation is less pronounced in districts that are "irregularly" shaped, i.e., not contained within a small area. Drawing on these observations, a policy paper by Brill (2019) proposed, "Another option to disrupt the segregating influence of catchment areas is to gerrymander them in reverse – that is, to draw catchment area lines to purposely include a diverse population." At the same time, such drastic measures can come at a substantial cost (Castro et al. 2024), even for those groups that are meant to benefit from them: for example, some disadvantaged households may experience significant increases in travel time as a result of rezoning. Formulating and solving a fair districting problem, such as the one in this paper, can help quantify the cost of fairness and contribute to a more informed policy.

In addition to group fairness, we also impose the condition of *individual fairness*, meaning that the assignment of an individual to a facility depends only on the geographical location of the individual, and not on the demographic group to which the individual belongs. It is well-known that these attributes are correlated (Bayer et al. 2004), but explicit dependence of the decision on the demographic group is perceived as unfair and has been successfully challenged in court (U.S. Supreme Court 2007, 2023). In our setting, individual fairness means that we have a single partition that is universally applied to the entire population.

Individual and group fairness have been extensively studied in very general contexts, and the distinction between them can be made more precise as follows. Suppose that we randomly select

an individual from the population. The individual is described by attributes (X, Z). In our setting, X represents the geographical location of the individual (a point on the map), but in the more general literature it may model other information; for example, Zhu & Ryzhov (2025a) considers a hiring problem where X is a vector describing the qualifications of a job applicant. The *protected attribute* Z is a categorical variable indicating the demographic or socioeconomic group to which the individual belongs. We act upon the individual with some decision Y (in our setting, an assignment to a facility). Because Y depends on the observed random attributes, it is also a random variable. We then have the following definitions:

- Individual fairness (Barocas & Selbst 2016): Y is conditionally independent of Z given X.

- Group fairness (Dwork et al. 2012): Y is independent of Z.

One does not at all imply the other, and it is quite uncommon to achieve both individual and group fairness simultaneously in the same problem. In fact, an influential paper by Kleinberg et al. (2017) argues that this is impossible in general, with unavoidable tradeoffs between different definitions of fairness. In the specific context of partitioning, however, a striking result by Dvoretzky et al. (1951) proved the existence of partitions satisfying both criteria for any number of facilities and subpopulations. Unfortunately, the proof was non-constructive, and the question of how such partitions can be computed remained open. This question is fully resolved in our paper: we provide a complete characterization of partitions that not only satisfy both types of fairness, but are *optimal* with respect to expected cost.

Our approach is inspired by the theory of semidiscrete optimal transport (Hartmann & Schuhmacher 2020, Villani 2021), though our problem does not fully fit that framework due to the presence of the group fairness condition. We relax the partitioning problem to allow probabilistic assignments, so that a fixed location can be assigned to any facility randomly according to a distribution to be determined by the decision-maker. This version of the problem can be formulated as an infinite-dimensional linear program. Crucially, however, the dual of this problem can be written as a nonlinear, but finite-dimensional concave program, which can be tractably solved. The dual variables characterize the optimal assignment policy, which turns out to be deterministic; in other words, even with probabilistic assignments allowed, it is still optimal to assign each location to a single facility. The assignment rule has a geometric interpretation as a novel generalization of the additively weighted Voronoi diagram (Aurenhammer 1991), which is well-known to arise in classical

districting problems. In our context, each facility has multiple weights (one for each group) rather than just one, and these weights are mixed differently at each location depending on the relative incidence of each group there.

We then show how the optimal partition can be efficiently computed. Although the objective function of the finite-dimensional dual involves a difficult integral, it turns out that it can be tractably optimized using the method of stochastic approximation (Pasupathy & Kim 2011). We only require the ability to simulate individuals from the spatial distribution of any demographic group; in the case of population densities, these data are publicly available. With simulation, the problem can be solved using a very simple algorithm, which can be coded in several lines, and which provably converges at a near-optimal rate. Because the procedure only requires realizations of costs, we can actually solve the problem without knowing the closed form of the cost function. In fact, in our case study, we obtain costs from a database that uses real network data and accounts for traffic patterns in urban environments.

We demonstrate the efficacy and utility of our method on a subregion of Los Angeles with 78 school district offices and seven demographic groups for which census data were available. The case study starkly illustrates the costs and tradeoffs that arise when imposing demographic parity. We find that the imposition of fairness constraints significantly increases travel costs across the board for all groups, including disadvantaged ones. In addition, the presence of segregation can lead to undesirable geometry in the partition, e.g., discontinuities in the service zones. At the same time, our results also provide insight into how these negative impacts can be mitigated. For example, extreme outliers of the travel cost distribution can be reduced by optimizing expected *squared* travel time, which is trivial to do under our modeling and algorithmic framework. Furthermore, reducing the number of distinct groups in the model will improve costs across the board for *all* groups. Potentially, such an analysis could help policymakers decide the set of groups for which fair representation is to be enforced; in some cases, the cost of fairness may outweigh the benefits for certain subpopulations, while also increasing the burden on everyone else. The presence of such cases can also be interpreted as a measure of how much segregation is present in the region under consideration.

In summary, this paper makes the following contributions. 1) We formulate a new class of geographical partitioning problems with a heterogeneous population and demographic parity constraints, neither of which has previously been considered in computational geometry. 2) We prove

that the optimal solution is completely characterized by a finite set of weights, which comprise a novel generalization of the additively weighted Voronoi diagram. 3) We show how the optimal weights can be found using a stochastic approximation algorithm, which is trivial to implement, converges at a near-optimal rate, and requires only realizations from the cost function. In this way, we answer a long-standing open question about the computation of fair partitions. 4) We empirically evaluate our approach on a realistic setting, providing insight into the cost of fairness for each subgroup. Overall, our work combines concepts from algorithmic fairness and computational geometry in a novel way, and is of both theoretical and practical interest.

2 Literature review

Our work bridges three distinct streams of literature, including elements of algorithmic fairness, optimal transport, and geographical partitioning. Below, we position our work relative to each stream.

Algorithmic fairness. This literature, surveyed by del Barrio et al. (2020), has mainly grown out of the computer science community. Both individual (Joseph et al. 2016, Gupta & Kamble 2021) and group (Dwork et al. 2012, Kallus & Zhou 2021) fairness, in the sense of our paper, have attracted considerable attention. Only one of these definitions is generally adopted in any given study, due to the widespread belief that the differences between them are irreconcilable. In this respect, our work stands out as an example of a setting where two distinct definitions can be satisfied simultaneously.

Computer scientists and statisticians have also proposed other definitions specialized for prediction problems (Hardt et al. 2016, Chouldechova 2017), which are the main focus of machine learning. For example, given a limited sample, one may wish to balance prediction error (Calders & Verwer 2010), or reduce prediction bias (Zink & Rose 2020), across multiple demographic groups. These studies do not explicitly model any decision problem, and thus do not apply to situations where the goal is to design a fair *decision* rather than a fair prediction. See also Zhu & Ryzhov (2025a) for an in-depth example of a problem where fair decisions are very different from fair predictions.

Optimal transport. Recent work (Hartmann & Schuhmacher 2020) has shown that optimal transport theory (Peyré & Cuturi 2019) offers a useful lens through which to approach geographical partitioning problems. Semidiscrete optimal transport is an infinite-dimensional linear program

with a finite-dimensional dual (Carlsson et al. 2016), which can be solved efficiently using the method of stochastic approximation (Genevay et al. 2016), a classical technique from the simulation literature (Pasupathy & Kim 2011).

Recently, Johndrow & Lum (2019) and Chzhen et al. (2020) leveraged optimal transport theory to find policies satisfying group fairness conditions, by connecting the latter to the Wasserstein barycenter problem (Agueh & Carlier 2011). This approach is limited to quadratic cost functions, and also allows the decision to depend directly on the demographic group, thus violating individual fairness. In fact, when both types of fairness are considered, the problem is no longer reducible to classical optimal transport. However, we show that infinite-dimensional duality still applies and provides a roadmap for efficient computation.

Geographical partitioning. The field of computational geometry (Carlsson & Devulapalli 2013) has extensively studied additively weighted Voronoi diagrams (Aurenhammer & Klein 2000), a class of geometric structures where the assignment of an individual to a facility is made by minimizing the sum of travel distance and a facility-specific "weight." Applications include the design of delivery (Haugland et al. 2007) and political (Ricca et al. 2008) districts, as well as various problems in facility logistics (Yushimito et al. 2012). Most of these papers, like ours, view partitioning as a central planning problem, with Carlsson et al. (2024) being one exception where both centralized and decentralized behavior is considered. While this literature has examined some notions of fairness, mainly in the form of load-balancing constraints between facilities (Aronov et al. 2009), our paper is the first to explicitly model multiple demographic groups with different population densities. As a result, we discover a novel generalization of the Voronoi diagram in which each facility uses a location-dependent mixture of weights.

There is also a mature literature on districting that uses mixed-integer programming models; examples include Swamy et al. (2023), Shahmizad & Buchanan (2025), Ozel et al. (2025) and other references cited therein. Various measures of fair representation can be incorporated into the constraints of these models (Arredondo et al. 2021). This work is primarily application-driven. While our approach is more stylized, it enables us to derive a clean analytical characterization of partitions satisfying both individual and group fairness, providing an understanding of how group fairness impacts the geometric structure of the optimal partition.

3   Problem statement

Let $\mathcal{X}$ be a compact subset of $\mathbb{R}^d$ representing the geographical region under consideration (in most practical applications, we will have $d = 2$). Let X be a random variable following a probability density f supported on $\mathcal{X}$. We may think of f as the population density and X as the geographical location of a randomly selected individual from the population.

Let Z be a discrete random variable taking values in $\{1, ..., M\}$ with $q_z = P(Z = z)$. Each z represents a different subgroup of the population, with $f_z(x) = P(X \in dx \mid Z = z)$ denoting the conditional density for that group. In this paper, q_z and f_z are all known; in practice, one would extract them from census data.

Suppose that there are K facilities located at points $x_1, ..., x_K \in \mathcal{X}$. Let $c(x, k)$ be the cost of assigning an individual at location x to the kth facility. A natural and widely used cost function is the Euclidean distance $c(x, k) = \|x - x_k\|_2$, but our analysis is not limited to this choice. We only require the condition that the pairwise difference $c(X, j) - c(X, k)$, for any $j \neq k$, is absolutely continuous with respect to Lebesgue measure.

We formulate the problem

$$\min_{\mathcal{A}_1, ..., \mathcal{A}_K} \sum_k \mathbb{E}\left(c(X, k) \mathbf{1}_{\{X \in \mathcal{A}_k\}}\right) \tag{1}$$

subject to

$$P(X \in \mathcal{A}_j \cap \mathcal{A}_k) = 0, \quad j \neq k, \tag{2}$$

$$\bigcup_k \mathcal{A}_k = \mathcal{X}, \tag{3}$$

$$P(X \in \mathcal{A}_k \mid Z = z) = P(X \in \mathcal{A}_k \mid Z = z'), \quad z, z' \in \{1, ..., M\}, \quad k = 1, ..., K. \tag{4}$$

This problem partitions $\mathcal{X}$ into sets $\mathcal{A}_k$. Constraints (2) ensure that the sets are disjoint except possibly on a boundary that has measure zero. Constraint (3) ensures that the sets comprise a partition. An individual appearing at a location $x \in \mathcal{A}_k$ is presumed to be assigned to facility k, incurring cost $c(x, k)$. The objective (1) thus chooses the partition to minimize the expected cost incurred by a randomly selected individual. Thus far, these elements of the problem are typical of geographical partitioning problems, e.g., those in Carlsson et al. (2016).

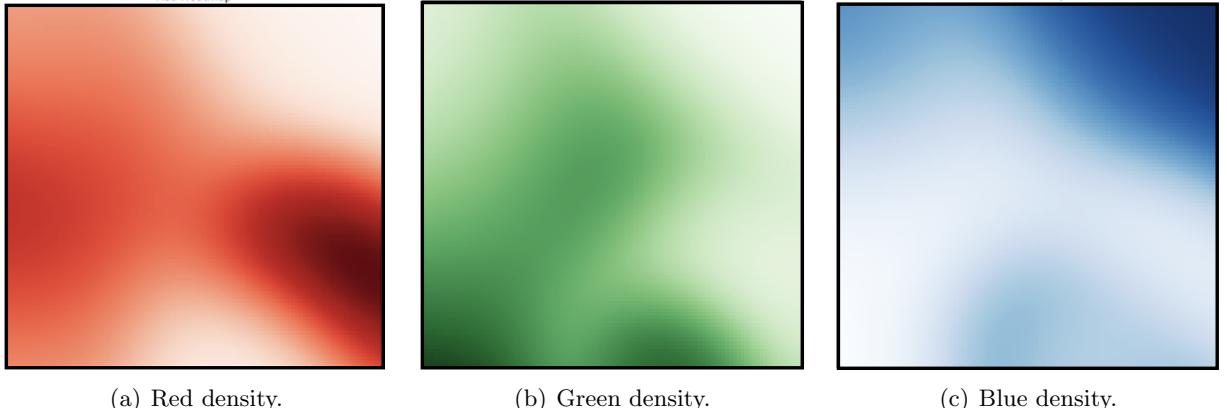In marked contrast with those problems, however, we also include the *group fairness* constraints

(a) Red density.   (b) Green density.   (c) Blue density.

Figure 1: Illustrative example with three subpopulations.

(4), which ensure that the region served by facility k is probabilistically independent of Z. In other words, each facility must serve the same proportion of each subgroup, or, to put it yet another way, the demographics of the individuals served by facility k should match those of the population as a whole. As discussed in Section 2, group fairness has been studied in a variety of settings, such as prediction (Chzhen et al. 2020) and hiring (Zhu & Ryzhov 2025a). The critical distinction here is that (1)-(4) also imposes *individual fairness*, that is, we design only a single partition that is applied to all demographic groups. Another way to say this is that the assignment of an individual is based only on the individual's location X and not their demographic Z, whereas virtually all existing papers on group fairness would allow the decision to explicitly depend on both.

In the context of geographical partitioning, it is possible to simultaneously achieve both types of fairness. Dvoretzky et al. (1951) proved that the system (2)-(4) is feasible, though this paper did not show how the partition could be computed, and indeed it did not explicitly model any objective function. In our framework, the objective (1) introduces the notion of an *optimal* fair partition. In Sections 4-5, we will show how such a partition can be characterized and computed. First, however, we present a simple example to illustrate the concept.

Figure 1 shows an instance where the geographical region $\mathcal{X} = [0,1]^2$ contains three subpopulations. The population densities f_z, for $z \in \{1, 2, 3\}$, are shown separately. We see that the "red" group is mainly concentrated in the bottom-right portion of the unit square, with some presence near the left edge; the "green" group mainly resides along the bottom edge, with some presence in the middle; and the "blue" group is heavily concentrated in the top-right. Each subpopulation has some presence in every part of the map, but overall the region is quite segregated.

In Figure 2, we superimpose the three densities onto the same plots and consider the partitioning problem with four facilities represented by white squares. Figure 2(a) shows a partition in which every individual is assigned to the closest facility, i.e., the facility with the smallest Euclidean distance from their location. This type of partition is very well-known in the literature under the name "Voronoi diagram." Individual fairness is clearly satisfied, because assignment is a deterministic function of location and has no additional dependence on demographic. However, group fairness is not satisfied: for example, the blue group is over-represented at facility 1 and under-represented at facility 4.

On the other hand, Figure 2(b) shows the *optimal fair partition* for this instance. The objective is still to minimize expected travel distance (measured using the Euclidean norm), but we now impose the group fairness constraints (4). The region served by facility z is now labeled. We see that the regions have become non-contiguous: for example, the top-right corner is now divided between facilities 1, 3, and 4, as the latter two have no other way to achieve representation for the blue group. Such geometry is likely undesirable in practice, and, in our case study in Section 6, we discuss some ways to reduce these effects by changing the cost function c. However, if the populations are already very heavily segregated, even an optimal partition will unavoidably have disconnected regions. The most that can be guaranteed is that the enclaves and exclaves will be chosen in a way that minimizes the overall cost at the population level.

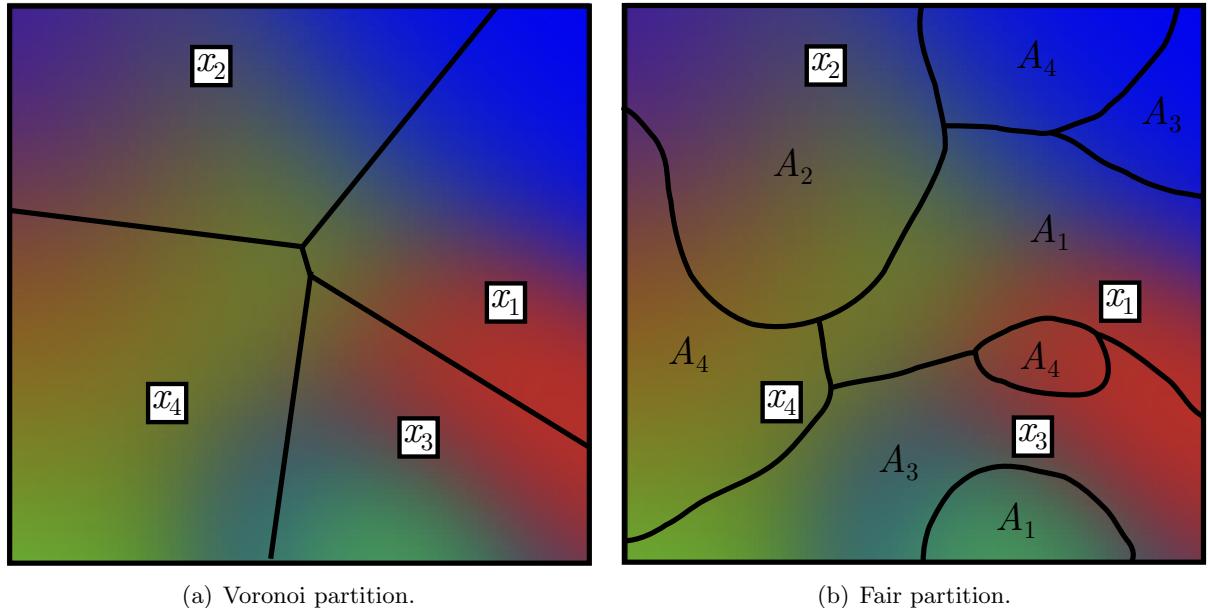(a) Voronoi partition.   (b) Fair partition.

Figure 2: Illustrative example with two partitions.

If we were to adopt the fair partition in this example, some individuals from the blue group would experience a significant increase in travel cost. At the same time, if the blue group were disadvantaged to begin with, the more distant facilities might be preferable in other ways (e.g., because those schools have more resources). Furthermore, whatever the resulting travel cost might be, the partition in Figure 2(b) is optimal, and any attempt to make the borders "nicer" would inevitably make the objective value even worse. One could also make the argument (though this issue is outside the scope of this paper) that group fairness will eventually change the population densities themselves in a manner that reduces segregation. Lastly, a comparison of the two partitions in Figure 2 could allow policymakers to quantify the cost of eliminating segregation, regardless of whether the fair districting plan is implemented. In other words, the study of (1)-(4) can be informative and insightful even when not interpreted as a direct "call to action."

4 Optimal transport formulation and solution

To solve (1)-(4), we relax the problem to allow probabilistic assignments, analogously to the well-known Kantorovich formulation of semidiscrete optimal transport (Villani 2021). Let Y be a discrete random variable taking values in $\{1, ..., K\}$. The quantity $g(x, k) = P(Y = k \mid X = x)$ represents the conditional probability of assigning an individual to facility k, given that they appear at location x. In the original formulation, this probability would be 1 if $x \in \mathcal{A}_k$ and 0 otherwise, but this may no longer be the case in the relaxation. We will eventually show, however, that the *optimal* solution of the relaxed problem is also feasible for (1)-(4). This result is similar to the well-known equivalence of the Monge and Kantorovich formulations of optimal transport (Pratelli 2007), with the difference that neither of those formulations contains fairness constraints.

In Section 4.1, we formulate and solve the relaxed problem under the additional assumption that $P(Y = k) = p_k$ for some pre-specified vector p of probabilities. This is like modifying (4) to require $P(X \in \mathcal{A}_k \mid Z) = p_k$. Then, in Section 4.2, we show how to optimize over p. We break the problem down into these two parts for ease of understanding, and also because the case of fixed p may be of stand-alone interest in itself (for example, if the kth facility only has enough capacity to handle a proportion p_k of the population).

4.1 Solution with fixed region sizes

We formulate the problem

$$\min_{g} \sum_{k} \iint_{\mathcal{X}} \sum_{z} q_z c(x, k) g(x, k) f_z(x) dx \tag{5}$$

subject to $g \geq 0$ and

$$\sum_{k} g(x, k) = 1, \quad x \in \mathcal{X}, \tag{6}$$

$$\iint_{\mathcal{X}} g(x, k) f_z(x) dx = p_k, \quad k = 1, ..., K, \quad z = 1, ..., M. \tag{7}$$

The objective function in (5) is simply $\mathbb{E}(c(X, Y))$ written out by conditioning on Z. This is the same objective as (1), with the sole difference being that the assignment Y is no longer required to be a deterministic function of X. Constraints (6) require every location to be assigned to a facility, analogously to (3). Constraints (7) force Y to be probabilistically independent of Z, similarly to (4), but also impose the fixed pmf p on the marginal distribution of Y. Individual fairness is ensured by the fact that, for fixed x and k, the decision $g(x, k)$ has no dependence on Z.

Problem (5)-(7) resembles semidiscrete optimal transport (Hartmann & Schuhmacher 2020, Zhu & Ryzhov 2025b). The difference is that, in the classical semidiscrete formulation, the index z is not present at all, and we can then simply optimize over the joint density of (X, Y) rather than the conditional pmf of Y given X. The constraints in semidiscrete optimal transport fix the marginal distributions of both X and Y, without the additional complication of fairness.

These differences notwithstanding, problem (5)-(7) has one aspect in common with semidiscrete OT, namely, that it is an infinite-dimensional linear program. We may therefore derive its Kantorovich dual using the following result.

Theorem 4.1. *The Kantorovich dual of problem (5)-(7) is given by*

$$\sup_{\phi, \psi_z} \iint_{\mathcal{X}} \phi(x) dx + \sum_{z} \sum_{k} p_k \psi_z(k) \tag{8}$$

subject to the constraints

$$\phi(x) + \sum_{z} f_z(x) \psi_z(k) \leq c(x, k) \left(\sum_{z} q_z f_z(x) \right), \quad x \in \mathcal{X}, \quad k = 1, ..., K. \tag{9}$$

The proof of Theorem 4.1 is given in the Appendix. The dual variables $\phi(x)$ and $\psi_z(k)$ correspond to the primal constraints (6) and (7), respectively. Note that (8)-(9) is also an infinite-dimensional linear program, with ϕ being a functional variable. However, as we will now show, it is possible to reformulate the problem as a nonlinear, but finite-dimensional concave program.

We may optimize ϕ by taking

$$
\begin{aligned}
\phi(x) &= \min_k \left[c(x,k) \left(\sum_z q_z f_z(x) \right) - \sum_z f_z(x) \psi_z(k) \right] \\
&= \left(\sum_z q_z f_z(x) \right) \min_k \left[c(x,k) - \sum_z \frac{f_z(x)}{\left(\sum_{z'} q_{z'} f_{z'}(x) \right)} \psi_z(k) \right].
\end{aligned}
$$

Letting $w_{k,z} = \frac{\psi_z(k)}{q_z}$, we may rewrite (8)-(9) as the unconstrained (and finite-dimensional) optimization problem

$$
\max_{w_{k,z}} \iint_{\mathcal{X}} \left(\sum_z q_z f_z(x) \right) \min_k \left[c(x,k) - \sum_z \frac{q_z f_z(x)}{\left(\sum_{z'} q_{z'} f_{z'}(x) \right)} w_{k,z} \right] dx + \sum_z \sum_k p_k q_z w_{k,z}. \quad (10)
$$

Observe, furthermore, that

$$
\sum_z q_z f_z(x) = \sum_z P(X \in dx \mid Z = z) P(Z = z) = P(X \in dx),
$$

whereas

$$
\frac{q_z f_z(x)}{\sum_{z'} q_{z'} f_{z'}(x)} = \frac{P(X \in dx \mid Z = z) P(Z = z)}{P(X \in dx)} = P(Z = z \mid X = x).
$$

Consequently, (10) may be further rewritten as

$$
\max_{w_{k,z}} \mathbb{E} \left\{ \min_k \left[c(X,k) - \mathbb{E}(w_{k,Z} \mid X) \right] \right\} + \sum_z \sum_k p_k q_z w_{k,z}. \quad (11)
$$

Before proceeding further, let us contrast (11) with the well-known dual reformulation of semidiscrete OT (Genevay et al. 2016). In that problem, Z is not present (the population is homogeneous), so the primal LP is essentially a special case of (5)-(7) with $M = 1$. In that case, (11) reduces to

$$
\max_{w_k} \mathbb{E} \left(\min_k c(X,k) - w_k \right) + p^\top w, \quad (12)
$$

that is, for each facility k there is a single, fixed weight w_k. In the presence of group fairness constraints, however, the kth facility now has multiple weights $w_{k,z}$ corresponding to different subgroups, and the objective function calculates a mixture of these weights according to the conditional

distribution of Z given X. However, since the conditional expectation is a linear function of the weights, the function

$$F(w, x) = \min_k [c(x, k) - \mathbb{E}(w_{k,Z} \mid X = x)],$$

viewed as a function of w, is a minimum of a finite number of linear functions and therefore concave. Since expectations preserve concavity, (11) is an unconstrained concave maximization problem, and therefore tractable in principle.

We can further characterize the optimal solution w^* of (11). Since $\mathcal{X}$ is closed and bounded, we may differentiate under the integral sign and obtain

$$\nabla_w \mathbb{E}[F(w, X)] = \mathbb{E}[\nabla_w F(w, X)].$$

For fixed x and k, z,

$$\frac{\partial}{\partial w_{k,z}} F(w, x) = -P(Z = z \mid X = x) \cdot 1_{\left\{k = \arg\min_j c(x,j) - \mathbb{E}(w_{j,Z} \mid X = x)\right\}}.$$

Therefore,

$$
\begin{aligned}
\frac{\partial}{\partial w_{k,z}} \mathbb{E}[F(w, X)] &= -\mathbb{E}\left[P(Z = z \mid X) \cdot 1_{\left\{k = \arg\min_j c(X,j) - \mathbb{E}(w_{j,Z} \mid X)\right\}}\right] \\
&= -\mathbb{E}\left[1_{\{Z=z\}} \cdot 1_{\left\{k = \arg\min_j c(X,j) - \mathbb{E}(w_{j,Z} \mid X)\right\}}\right],
\end{aligned}
\tag{13}
$$

where (13) follows by the projection property of conditional expectations (Cinlar 2011, ch. IV, sec. 1, eq. (1.4)). Therefore, w^* solves the system

$$\mathbb{E}\left[1_{\{Z=z\}} \cdot 1_{\left\{k = \arg\min_j c(X,j) - \mathbb{E}(w_{j,Z} \mid X)\right\}}\right] = p_k q_z, \quad k = 1, ..., K, \quad z = 1, ..., M.
\tag{14}$$

From (11), we can also characterize the optimal primal solution as a function of the optimal dual variables. The following result shows how to construct Y to optimize the relaxation (5)-(7).

Proposition 4.1. *Let w^* denote the optimal solution of (11), and define*

$$Y^* = \arg\min_j c(X, j) - \mathbb{E}\left(w_{j,Z}^* \mid X\right),$$

with ties broken arbitrarily. Then, the joint distribution of (X, Y^*) optimally solves (5)-(7).

Proof: By definition of Y^*, (14) is exactly $P(Z = z, Y^* = k) = p_k P(Z = z)$, whence $P(Y^* = k \mid Z = z) = p_k$ for all z. This choice of Y is therefore feasible for (6)-(7).

By weak duality,

$$\mathbb{E}\left\{\min_k \left[c\left(X,k\right) - \mathbb{E}\left(w_{k,Z}^* \mid X\right)\right]\right\} + \sum_z \sum_k q_z p_k w_{k,z}^* \leq \min_g \sum_k \mathbb{E}\left(c\left(X,k\right) 1_{\{Y=k\}}\right) dx,$$

where g is the conditional pmf of Y given X, satisfying (6)-(7) as before. We then write

$$
\begin{aligned}
\mathbb{E}\sum_k c\left(X,k\right) 1_{\{Y^*=k\}} &= \mathbb{E}\sum_k \left(c\left(X,k\right) - \mathbb{E}\left(w_{k,Z}^* \mid X\right)\right) 1_{\{Y^*=k\}} + \mathbb{E}\sum_k \mathbb{E}\left(w_{k,Z}^* \mid X\right) 1_{\{Y^*=k\}} \\
&= \mathbb{E}\left\{\min_k \left[c\left(X,k\right) - \mathbb{E}\left(w_{k,Z}^* \mid X\right)\right]\right\} + \sum_k \mathbb{E}\left(w_{k,Z}^* 1_{\{Y^*=k\}}\right) \quad (15) \\
&= \mathbb{E}\left\{\min_k \left[c\left(X,k\right) - \mathbb{E}\left(w_{k,Z}^* \mid X\right)\right]\right\} + \sum_k \mathbb{E}\left(w_{k,Z}^*\right) P\left(Y^* = k\right) \quad (16) \\
&= \mathbb{E}\left\{\min_k \left[c\left(X,k\right) - \mathbb{E}\left(w_{k,Z}^* \mid X\right)\right]\right\} + \sum_k \sum_z p_k q_z w_{k,z}^*.
\end{aligned}
$$

where (15) again follows by the projection property of conditional expectations (since Y^* is a deterministic function of X), whereas (16) uses the independence of Z and Y^*. The desired result follows. $\square$

The optimal Y^* from Proposition 4.1 then induces the partition

$$\mathcal{A}_k = \left\{x \in \mathcal{X} : k = \arg\min_j c\left(x,j\right) - \mathbb{E}\left(w_{j,Z}^* \mid X = x\right)\right\}, \tag{17}$$

which is feasible for the original problem (1)-(4) under some mild conditions on the distribution of X. Specifically, it suffices for the pairwise differences $c\left(X,j\right) - c\left(X,k\right)$ to have a density for any $j \neq k$.

Once again, it is interesting to compare (17) with the partition obtained in the standard semidiscrete OT problem, where group fairness constraints are not present. Recall that this standard problem has the dual objective (12), where there is no Z variable and only one weight for each k value. Letting w_k^* be the optimal weights for that problem, (12) induces a partition where an individual appearing at location x is assigned to the facility corresponding to $\arg\min_j c\left(x,j\right) - w_j^*$. This geometric structure, where assignments are made by minimizing the sum of travel distance and a facility-specific weight, is precisely the well-known additively weighted Voronoi diagram (Aurenhammer 1991). In our setting, (11) induces the partition (17), which can be viewed as a novel generalization of this concept. The imposition of group fairness constraints leads to a diagram in which each facility has multiple weights, rather than just one, and the assignment uses a mixture of these weights that depends on the location of the individual. In other words, if a subgroup z is

14

more prevalent at location x, then the corresponding weights $w_{j,z}^*$ play a greater role in determining the facility serving that location.

4.2 Solution with optimal region sizes

We will now show how the vector p may be eliminated entirely, yielding an optimal solution to (1)-(4) under the best possible marginal distribution of Y. Let $w^*(p)$ denote the optimal weights for problem (11) with a fixed p, and let

$$V^*(p) = \mathbb{E}\left\{ \min_k \left[c(X, k) - \mathbb{E}\left(w_{k,Z}^*(p) \mid X \right) \right] \right\} + \sum_z \sum_k p_k q_z w_{k,z}^*(p). \tag{18}$$

By the envelope theorem,

$$\frac{\partial}{\partial p_k} V^*(p) = \sum_z q_z w_{k,z}^*(p),$$

because the only direct dependence of (18) on p occurs in the last term.

Consider the outer optimization problem

$$\min_p V^*(p)$$

subject to $p_k \geq 0$ and $\sum_k p_k = 1$. Let ν_k and λ be the Lagrange multipliers of the inequality and equality constraints, respectively. Then, the optimal p^*, λ^*, ν^* satisfy the first-order optimality conditions

$$\sum_z q_z w_{k,z}^*(p^*) + \lambda^* - \nu_k^* = 0, \qquad k = 1, ..., K, \tag{19}$$

$$\nu_k^* p_k^* = 0, \qquad k = 1, ..., K, \tag{20}$$

$$\nu_k^* \geq 0, \qquad k = 1, ..., K. \tag{21}$$

Adding a constant λ^* to all the weights will not affect the partition, so we may assume $\lambda^* = 0$ without loss of generality. Then, the complementary slackness conditions imply that

$$\sum_z q_z w_{k,z}^*(p^*) = 0 \qquad \Leftrightarrow \qquad p_k^* > 0,$$

$$\sum_z q_z w_{k,z}^*(p^*) = \nu_k^* \qquad \Leftrightarrow \qquad p_k^* = 0.$$

For k satisfying $p_k^* = 0$, define the modified weights $\bar{w}_{k,z} = w_{k,z}^*(p^*) - \nu_k^*$. Then, clearly $\sum_z q_z \bar{w}_{k,z} = 0$. To show that the modified weights do not change the partition, it is sufficient to show that the

set $\mathcal{A}_k$ induced by $w_{k,z}^* (p^*)$ according to (17) remains unchanged by the modification. Since $p_k^* = 0$, we have

$$c (x, k) - \mathbb{E} \left(w_{k,Z}^* (p^*) \mid X = x \right) \geq \min_{j \neq k : p_j^* > 0} c (x, j) - \mathbb{E} \left(w_{j,Z}^* (p^*) \mid X = x \right), \qquad x \in \mathcal{X}, \quad (22)$$

that is, for any x there is always one facility that is preferable to k. Suppose now that, under the modified weights, (22) is violated: that is, there exists x_0 such that

$$c (x_0, k) - \mathbb{E} \left(\bar{w}_{k,Z} \mid X = x_0 \right) < \min_{j \neq k : p_j^* > 0} c (x_0, j) - \mathbb{E} \left(w_{j,Z}^* (p^*) \mid X = x_0 \right).$$

However,

$$
\begin{aligned}
c (x_0, k) - \mathbb{E} \left(\bar{w}_{k,Z} \mid X = x_0 \right) &= c (x_0, k) - \mathbb{E} \left(w_{k,Z}^* (p^*) \mid X = x_0 \right) + \nu_k^* \\
&\geq c (x_0, k) - \mathbb{E} \left(w_{k,Z}^* (p^*) \mid X = x_0 \right),
\end{aligned}
$$

which leads to a contradiction with (22). Thus, the modified weights do not change the original partition; the objective value is also unchanged since it is determined only by the non-empty regions. Therefore, the optimality conditions (19)-(21) of the outer problem can be rewritten more succinctly as

$$\sum_z q_z w_{k,z}^* (p^*) = 0, \quad k = 1, ..., K,$$

that is, $\mathbb{E} \left(w_{k,Z}^* \right) = 0$ for all k.

5 Characterization and computation of optimal partition

We may now combine the results of Sections 4.1-4.2 and give a complete characterization of the optimal partition. Our goal is to find a solution (w^*, p^*) of the system of nonlinear equations

$$\mathbb{E} \left[1_{\{Z=z\}} \cdot 1_{\left\{ k = \arg\min_j c(X,j) - \mathbb{E} \left(w_{j,Z} \mid X \right) \right\}} \right] = p_k q_z, \quad k = 1, ..., K, \quad z = 1, ..., M, \quad (23)$$

$$\sum_z q_z w_{k,z} = 0, \quad k = 1, ..., K. \quad (24)$$

Note that the left-hand side of (23) is $P \left(Z = z, Y = k \right)$, so this constraint implies $p_k \geq 0$, $\sum_k p_k = 1$ and we may omit those conditions.

It can be readily seen that (23)-(24) are precisely the KKT conditions of the constrained concave program

$$\max_{w_{k,z}} \mathbb{E} \left\{ \min_k \left[c\left(X, k\right) - \mathbb{E}\left(w_{k,Z} \mid X\right) \right] \right\} \tag{25}$$

subject to the linear equalities (24). Here, the optimal region sizes p_k^* become the Lagrange multipliers of the equality constraints. This problem trivially satisfies Slater's condition, so (23)-(24) are necessary and sufficient for optimality. Even if there are multiple (w^*, p^*) satisfying (23)-(24), each solution achieves the same value of (25) as well as $V^*\left(p^*\right)$.

We may eliminate the constraints entirely by using the transformation

$$w_{k,z} = v_{k,z} - \frac{\sum_{z'} q_{z'} v_{k,z'}}{q^\top q} q_z = v_{k,z} - \frac{\mathbb{E}\left(v_{k,Z}\right)}{q^\top q} q_z, \tag{26}$$

which projects any arbitrary $(v_{k,1}, ..., v_{k,M})$ onto the orthogonal complement of q. This yields the unconstrained optimization problem

$$\max_{v_{k,z}} \mathbb{E} \left\{ \min_k \left[c\left(X, k\right) - \mathbb{E}\left(v_{k,Z} \mid X\right) + \frac{\mathbb{E}\left(q_Z \mid X\right)}{q^\top q} \mathbb{E}\left(v_{k,Z}\right) \right] \right\}. \tag{27}$$

Since (26) is a linear function of v, the objective function in (27) is still concave. Note that $\mathbb{E}\left(q_Z \mid X\right)$ does not depend on v. Therefore, for fixed k, z, and x,

$$\frac{\partial}{\partial v_{k,z}} \mathbb{E}\left(w_{k,Z} \mid X = x\right) = P\left(Z = z \mid X = x\right) - \frac{\mathbb{E}\left(q_Z \mid X = x\right)}{q^\top q} q_z.$$

Applying the projection property of conditional expectation once more, the optimality conditions of (27) are

$$\mathbb{E} \left\{ 1_{\left\{ k = \arg\min_j c(X,j) - \mathbb{E}\left(v_{j,Z} \mid X\right) + \frac{\mathbb{E}(q_Z \mid X)}{q^\top q} \mathbb{E}\left(v_{k,Z}\right) \right\}} \left(1_{\{Z=z\}} - \frac{q_Z}{q^\top q} q_z \right) \right\} = 0. \tag{28}$$

for $k = 1, ..., K$ and $z = 1, ..., M$.

The system (27) may be approached as a stochastic root-finding problem (Pasupathy & Kim 2011). Because the expectation in (28) is difficult to compute, we may instead approximate it using simulation, in a procedure known as stochastic approximation (Kushner & Yin 2003). Given an approximation v^n of the solution, we observe a sample $\left(X^{n+1}, Z^{n+1}\right)$ from the underlying joint distribution, and evaluate the random variable inside the expectation in (28) using these values. On average, this random variable will point us toward the root. We then compute v^{n+1} by adjusting v^n in the indicated (noisy) direction, using a stepsize to smooth out noise. Algorithm 1 formally states the procedure.

17

Algorithm 1 Stochastic approximation procedure for solving (27).

Step 0: Initialize $n = 0$ and $v_{k,z}^0$ for $k = 1, ..., K$, $z = 1, ..., M$. Choose a suitable deterministic stepsize sequence $\{\alpha_n\}_{n=0}^\infty$.

Step 1: Observe $\left(X^{n+1}, Z^{n+1}\right)$ and iterate

$$v_{k,z}^{n+1} = v_{k,z}^n + \alpha_n \cdot 1_{\left\{k = \arg\min_j c(X^{n+1}, j) - \mathbb{E}\left(w_{j,Z}^n | X = X^{n+1}\right)\right\}} \left(1_{\{Z^{n+1} = z\}} - \frac{q_{Z^{n+1}}}{q^\top q} q_z\right), \quad (29)$$

where w^n is computed by plugging v^n into (26).

Step 2: Increment n by 1 and return to Step 1.

Since the random variable in (29) is bounded, the convergence of $\{v^n\}$ to an optimum of (27) easily follows by classical SA theory. It is possible to characterize the convergence rate in greater depth. The following theorem (proved in the Appendix) derives a bound that matches the near-optimal rates in Bach & Moulines (2011), but uses weaker conditions. More precisely, Thms. 4-7 in Bach & Moulines (2011) require the stochastic gradient to be Lipschitz continuous, which does not hold in our setting because (29) uses indicator functions. We drop this requirement, but obtain the same guarantee. It is worth noting that the rate bound is obtained, not on the sequence $\{v^n\}$ directly, but on the *average* of these iterates. This technique, known as Polyak averaging, has been known to speed up convergence since Polyak & Juditsky (1992). From an implementation perspective, it is quite simple to incorporate averaging into Algorithm 1.

Theorem 5.1. *Let $v^0 \in \mathbb{R}^\ell$ for arbitrary dimension ℓ. Let $h : \mathbb{R}^\ell \to \mathbb{R}$ be a concave, differentiable function with $v^* \in \arg\max_v h(v)$. Let $\{\zeta^n\}_{n=1}^\infty$ be a sequence of random functions mapping $\mathbb{R}^\ell$ into $\mathbb{R}^\ell$ satisfying $\mathbb{E}(\zeta^n(v)) = \nabla_v h(v)$ and $P(|\zeta^n(v)| \le C) = 1$ for some constant $C > 0$. Define*

$$v^{n+1} = v_n + \alpha_n \zeta^n(v^n), \qquad n \ge 0,$$

with $\alpha_n = \frac{\alpha}{\sqrt{n+1}}$, and let $\bar{v}^n = \frac{1}{n} \sum_{m=1}^n v^m$ be the average of the iterates.

Then, there exists a constant $D > 0$, depending only on α, C, and $\|v_0\|_2$, such that

$$h(v^*) - \mathbb{E}(h(\bar{v}^n)) \le D \cdot \frac{\log n}{\sqrt{n}}.$$

6 Case study: Districting in LA County

In this section, we demonstrate the practicability of our approach on real population and demographic data. Our setting is the San Fernando Valley (SFV) region of Los Angeles County, a diverse area with significant demographic variation across neighborhoods. Some areas show high concentrations of specific demographic groups, a pattern typical of urban segregation documented in the literature (Reardon & Bischoff 2011).

The census distinguishes between seven[1] racial/ethnic categories: Multiracial (13.43% of the total population in the region), White (45.39%), Black (3.86%), American Indian (1.30%), Asian (11.41%), Pacific Islander (0.11%), and Other (24.50%). In Section 6.1, we apply our modeling and algorithmic framework to construct a partition that achieves fair representation for all seven categories, to demonstrate the capabilities of our approach. In Section 6.2, we repeat the same analysis with three groups: Black, Asian, and all the others combined into one. Our intention here is simply to demonstrate how the number of groups influences the costs for the population as a whole. For example, a comparison between the two versions of the case study will show that, when we require fair representation for all seven groups, costs increase for everyone across the board, including the disadvantaged populations whom the fair partition aims to benefit. This should be an important consideration for policymakers when deliberating the issue of fair representation.

6.1 Analysis for seven subgroups

We obtained latitude and longitude coordinates for school district administrative offices from the National Center for Education Statistics Education Demographic and Geographic Estimates Program geocoded data files (NCES 2022). From this dataset, we selected 78 administrative offices located within the SFV boundary using a San Fernando Valley shapefile obtained from ArcGIS Online ESRI (2025). We also incorporated detailed census tract population data from the American Community Survey for all seven demographic groups.

The cost $c(x, k)$ of traveling to district office k for a household located at coordinates x was computed using the Open Source Routing Machine (OSRM) with real road network data, providing realistic driving times and distances rather than simply Euclidean norms. This approach captures

---

[1]Unfortunately, "Hispanic" is not among the available options in the census, and therefore we were not able to include it as a distinct category in this study.

Race	Median (u)	Median (c)	Median (c-sq)	90th % (u)	90th % (c)	90th % (c-sq)
Multiracial	2034	3137	5394	3739	13536	10696
White	2217	3439	5955	4709	14578	12327
Black	2154	3439	5958	3739	18353	12510
American Indian	1853	2860	5363	3487	13536	10045
Asian	2213	3562	6058	4104	14319	12116
Pacific Islander	2068	4185	6104	3901	14590	11863
Other	1843	2991	5540	3455	14037	10045

Table 1: Travel distances (in meters) for unconstrained (u), equity-constrained (c), and equity-constrained with squared distance (c-sq) models, categorized by race.

the true cost of travel in an urban environment with complex street networks and traffic patterns. Using real costs does not create any additional challenges for our solution technique since Algorithm 1 only requires the ability to evaluate $c(X, k)$ for a realization X from the population. We ran our algorithm for two variants of the optimization problem: one where $c(x, k)$ was set equal to the real travel distance (which we also refer to as "regular distance"), and another one where it was set equal to the *squared* travel distance. In computational geometry, partitions based on squared distance are known as "power diagrams" and tend to create more compact and balanced regions, because they heavily penalize extreme deviations from district centers (Cohen-Addad et al. 2018).

Both variants include group fairness constraints. As a baseline, we also compute a standard Voronoi diagram where facility k serves the set $\{x : k = \arg\min_j c(x, j)\}$, that is, each household is assigned to the closest district office. This "unconstrained" partition will clearly not achieve demographic parity. Of course, in practice, the determination of school district boundaries involves other factors aside from simply travel distance. Nonetheless, Voronoi diagrams have been previously studied for this purpose (Pearce 2000, Gomilla & Arangala 2020), and because our focus here is on evaluating the price of fairness, they provide a natural basis for comparison. In these experiments, we do not fix the capacity of each facility (that is, we optimize over p_k), but if one wished to fix them in advance, this could be done using the techniques of Section 4.1.

Table 1 shows the mean and 90th percentile of the travel distance for each demographic group under each of the three models (for the model that optimizes squared cost, we still report statistics for the unsquared distance). Figure 6.1 further shows the empirical cumulative distribution functions of the travel distance for each demographic group/model combination. We see that the price of fairness is quite substantial, with all seven demographic groups experiencing significant increases in travel time relative to the unconstrained model. Optimizing the squared distance reduces out-
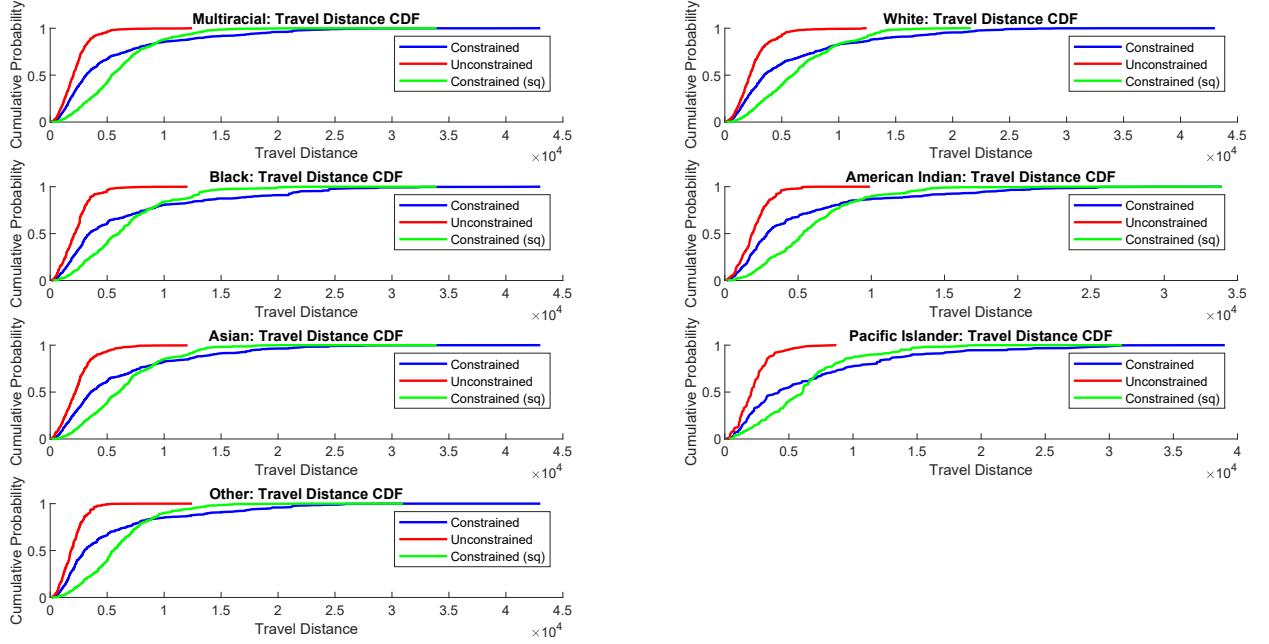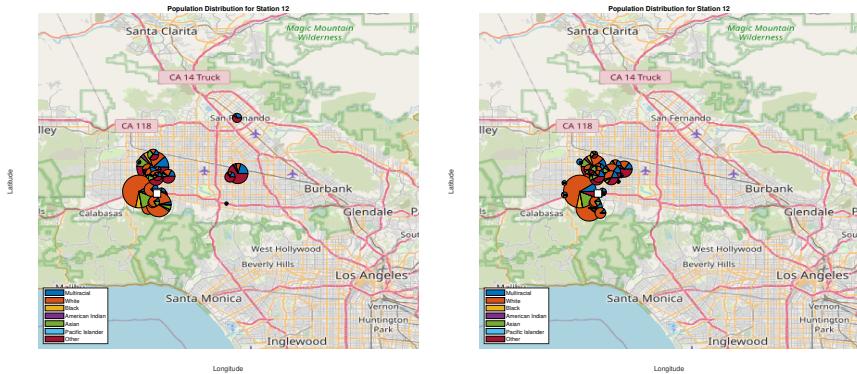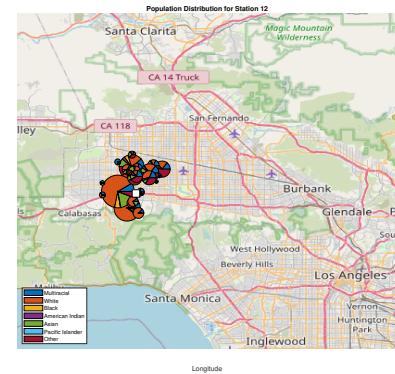
Figure 3: CDFs for travel distances for each demographic.

liers on both the low and high ends: the median travel time increases for each group relative to the standard equity-constrained model, with the benefit of reducing extremely high times.
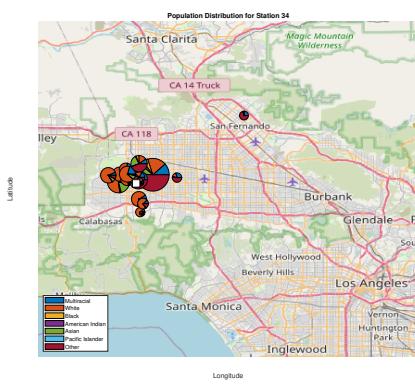
Additional insights can be obtained by examining individual district boundaries. Figure 4 shows four representative districts under both cost functions (a complete set of plots for all 78 facilities is included in the Appendix). The district offices are represented by white squares, and the census tracts assigned to the office in question are shown as pie charts whose size reflects the local population. Thus, for example, Facility 12 is located in the center-west of the map (along Route 101, east of Calabasas). Under regular distance, several census tracts from the other side of I-405 are assigned to this office, leading to high travel times for those households; under squared distance, these assignments are no longer made. Similar behavior can be observed for Facilities 34, 63, and 76: in all three cases, switching to squared distance eliminates the worst outliers, though it may not eliminate discontinuities completely. The optimal district sizes (that is, the optimal choices of p_k) need not the same for both cost functions. We will revisit these four examples in Section 6.2 when the entire case study is rerun for a reduced set of demographic groups.
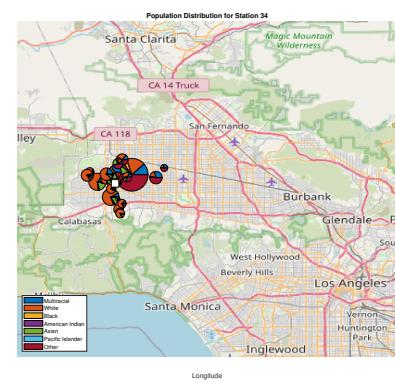
(a) Facility 12 (travel distance).
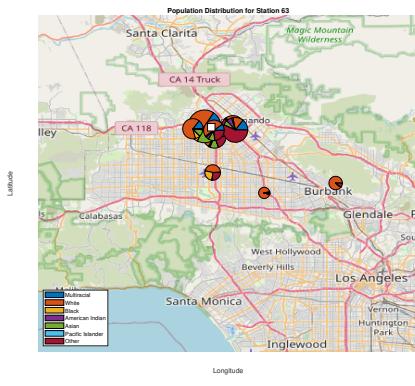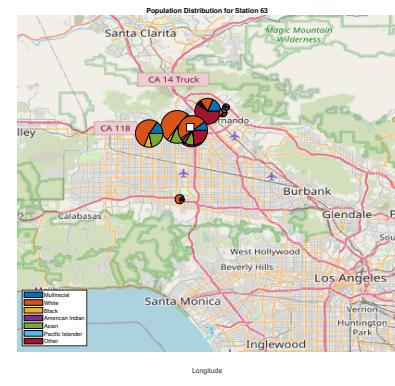
(b) Facility 12 (squared distance).

(c) Facility 34 (travel distance).
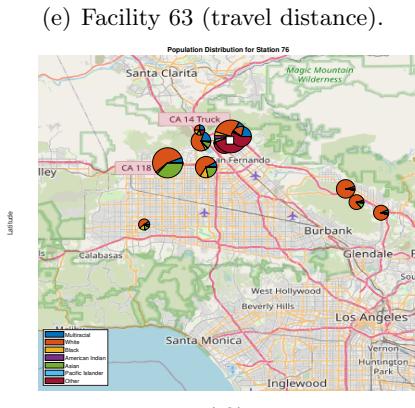
(d) Facility 34 (squared distance).
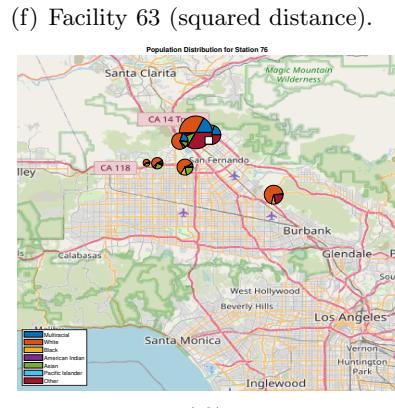
(e) Facility 63 (travel distance).

(f) Facility 63 (squared distance).

(g) Facility 76 (travel distance).

(h) Facility 76 (squared distance).

Figure 4: Four representative districts, drawn using two different cost functions (regular and squared distance).

It is also noteworthy that fair partitioning may result in the closure of school administrative offices. This does not happen for the unconstrained model, since for any facility, there is always some subset of the population that is closest to it. However, under equity constraints and regular cost, 7 out of 78 district offices receive no student assignments, and this number increases to 12 when we optimize the squared distance. Five offices are closed under both models. These situations can arise when a particular demographic group is hyper-concentrated in a small geographical area: if that area is very far away from a certain office, it may be more cost-effective to close that office entirely rather than shuttle individuals across the map to it.

While the results of this analysis may not be directly implementable, they can nonetheless be useful to inform discussion of equity in districting. First, it bears repeating that an equitable partition is shown to be *feasible*, which is not obvious without our analysis. Second, we are able to quantify the cost paid by each demographic group for fair representation. As Table 1 shows, the cost may even be higher for disadvantaged groups: for example, median travel times (under regular distance) increase by 54% for the white population, but by 59% for the Black population. This is a serious concern, and potentially a governing body might elect not to change district boundaries at all on such grounds. Third, if the authorities wished to adjust boundaries without fully adopting our approach, visuals such as those in Figure 4 may help guide these modifications by suggesting where representation of certain groups might be drawn from geographically. Finally, it is important to note that certain undesirable aspects of the boundaries (such as non-contiguity) are a consequence of a high level of segregation in the population. In a sense, the difficulty of creating an equitable partition is itself an illustration of how much segregation is already present.

6.2   Analysis for three subgroups

As mentioned previously, we also ran a version of the same case study in which only three demographic groups were considered: Black, Asian, and all the others combined. Thus, every facility must have 84.73% of its assigned population coming from this last group, but the precise breakdown within that group (White, Multiracial, Other and so on) is no longer constrained. The main purpose of these additional experiments is to show how the number of groups included in the model influences travel costs for all groups, whether explicitly included or not.

Table 2 presents the median and 90th percentile of the travel time for each group. Note that, while the *model* only distinguishes between three groups, we report results for *all seven* of the

Race	Median (u)	Median (c)	Median (c-sq)	90th % (u)	90th % (c)	90th % (c-sq)
Multiracial	2034	2473	3702	3739	6870	7783
White	2217	2547	3535	4709	7500	7958
Black	2154	2984	4261	3739	16592	11750
American Indian	1853	2241	3434	3487	5463	7490
Asian	2213	2943	4148	4104	14823	9963
Pacific Islander	2068	2683	3480	3901	8701	8042
Other	1843	2233	3606	3455	5390	7160

Table 2: Travel distances (in meters) for unconstrained (u), equity-constrained (c), and equity-constrained with squared distance (c-sq) models, categorized by race (three-group formulation).

groups from the original analysis. Once more, the baseline is a model with no fairness constraints (Voronoi diagram), so those statistics are the same as in Table 1.

When fairness constraints are included, the results are very different between Tables 1 and 2. Let us first consider the two groups (Black and Asian) that achieve fair representation under both versions of the case study. In Table 1 (seven groups), the median travel cost for the Black subpopulation increases by 59.6% relative to the baseline (under the regular cost function), but in Table 2, this increase is only 38.5%. Likewise, for the Asian subpopulation, the increase in the median is 60.9% with seven groups in the model, but only 32.9% with three groups.

For the remaining five groups, the difference is even greater. In Table 1, the median travel cost
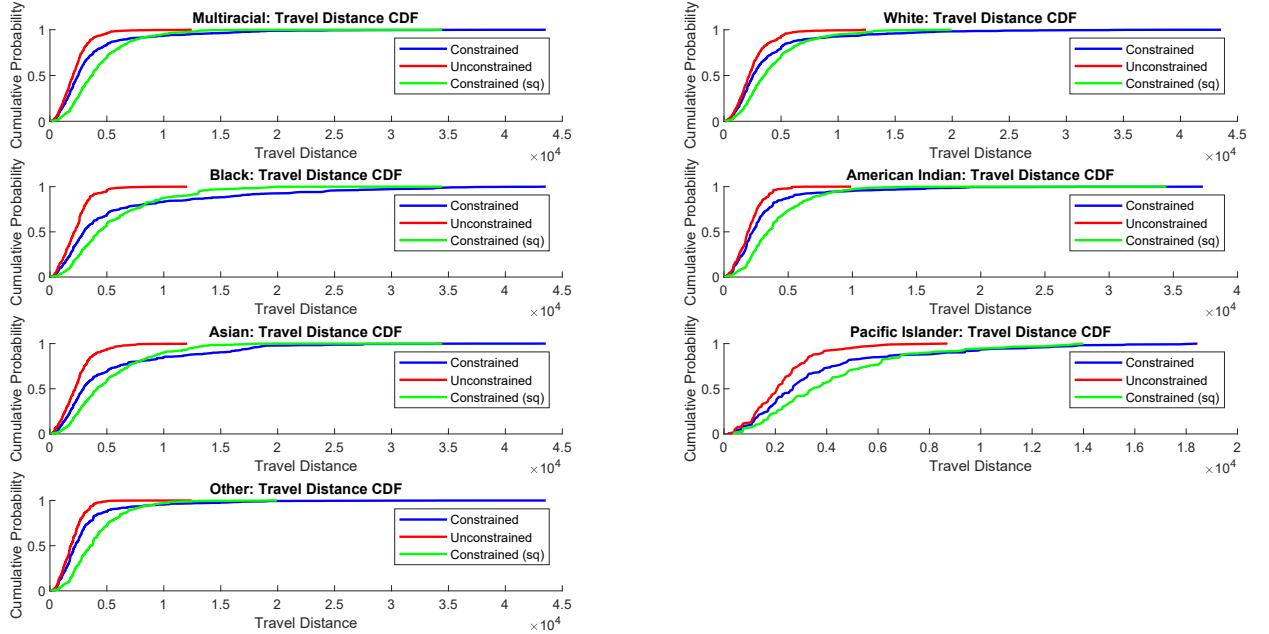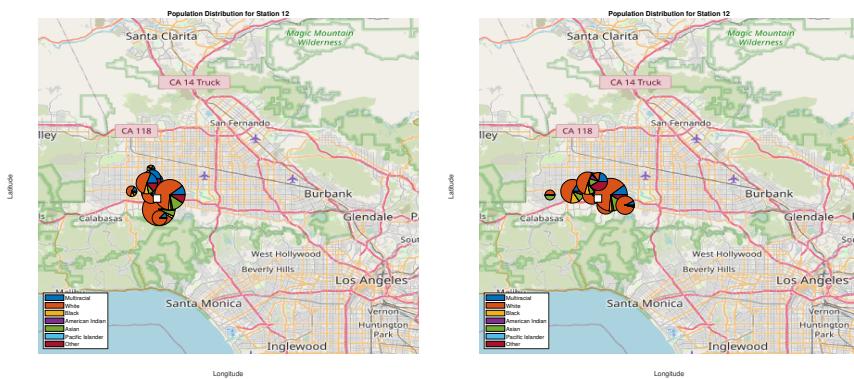
Figure 5: CDFs for travel distances for each demographic (three-group formulation).
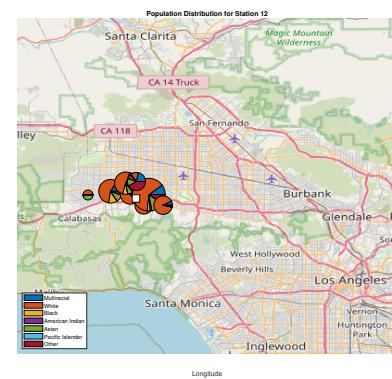
increases by over 54% for each of the five relative to the baseline model (again, under the regular cost function). In Table 2, the increase in the median is below 30% for all five groups. A more detailed comparison of the cdfs for various groups and models can be seen in Figure 6.2; in general, the distributions for the constrained models are closer to the baseline than they were in Figure 6.1.

Figure 6 revisits the four representative facilities from Figure 4 in the three-group context; a complete set of plots can be found in the Appendix. It is especially instructive to compare the district boundaries between the seven-group and three-group settings. For example, in the seven-group case (Figure 4), Facility 63 had three outlying census tracts under regular distance and one under squared distance. In the three-group case (Figure 6), the most distant outliers have disappeared and the census tracts are now closer to the office. Similarly, in the seven-group case, Facility 76 had several very distant outliers, far southwest and southeast of the district office, under regular distance; in the three-group case, only the southwestern outlier is still present, while the southeastern ones are no longer assigned to this facility. Under squared distance, the census tracts were closer to the facility to begin with, and again the most distant ones have been removed in the three-group case. In other words, running the model with three groups does not guarantee that the districts will become contiguous, but it significantly reduces the fragmentation of the boundaries. It is worth noting that district closures are also reduced: under three groups, 3 and 5 offices are closed, respectively, for regular and squared distance (in the seven-group case these numbers were 7 and 12). A single office (#16) is closed in all four versions of the case study.
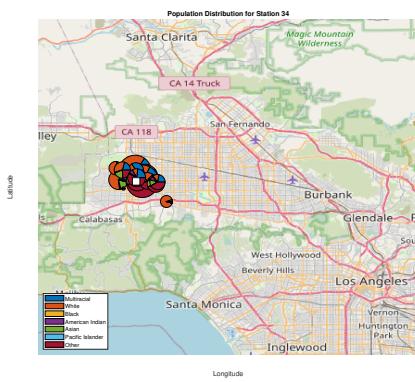
This discussion shows how our analysis can be used, not only to assess the cost of fairness in a general sense, but to make a more fine-grained determination of which subpopulations could be included in the model. For example, the Pacific Islander population is known to be disadvantaged in education (Pang et al. 2011). At the same time, when all seven groups are included in the model, Pacific Islanders suffer disproportionately from increased travel costs, with the median going up by 102% relative to the baseline (under unsquared costs). In the three-group version of the analysis, the increase for that subpopulation is only 29%. The precise considerations behind the decision to require fair representation for a particular group are far beyond the scope of our work, but it is conceivable that, in some cases, the cost of representation may be judged to be too prohibitive *for the disadvantaged group itself*. In such cases, reducing the total number of groups can also alleviate the burden on other groups that remain in the model.
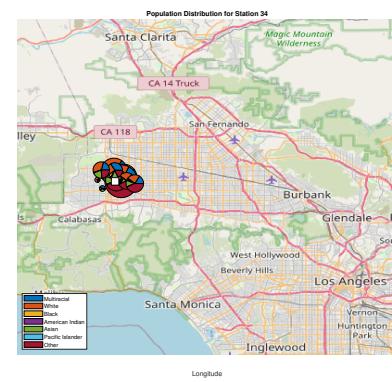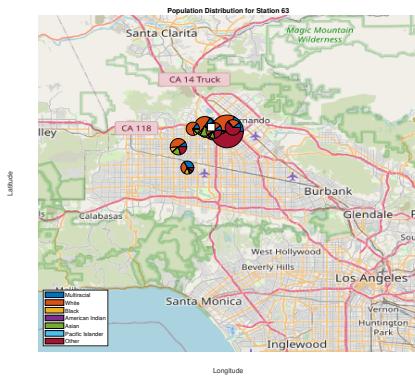
(a) Facility 12 (travel distance).

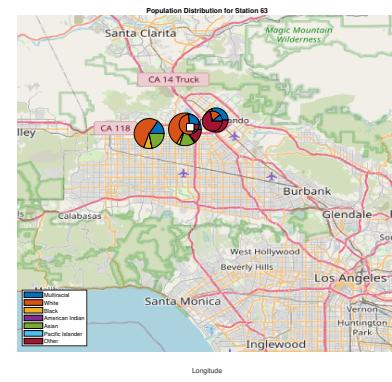(b) Facility 12 (squared distance).

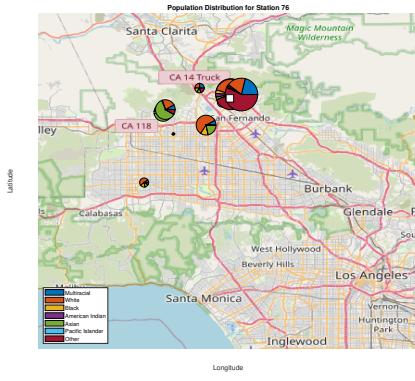(c) Facility 34 (travel distance).
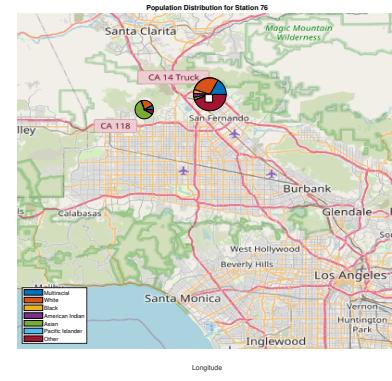

(d) Facility 34 (squared distance).

(e) Facility 63 (travel distance).

(f) Facility 63 (squared distance).

(g) Facility 76 (travel distance).

(h) Facility 76 (squared distance).

Figure 6: Four representative districts, drawn using two different cost functions (regular and squared distance).

7 Conclusion

We have introduced a new class of geographical partitioning problems in which the population is heterogeneous, and different subgroups are required to be represented fairly (i.e., proportionally to their incidence in the population as a whole) in each district. Our study is motivated by the phenomenon of socioeconomic segregation in school districts, where some demographic groups are over- or under-represented in certain areas. This problem has been a subject of discussion for many years among social scientists and policymakers, with some authors even recommending a kind of "reverse gerrymandering" as a means to achieve integration. Essentially, our work shows how this can be done in a principled and *optimal* way, constructing a fair partition that minimizes expected cost.

We have provided a complete mathematical characterization of this optimal fair partition, showing that it is a novel generalization of a classical geometric structure known as the additively weighted Voronoi diagram. Although the optimization problem we are solving is infinite-dimensional, it has a finite-dimensional dual, whose decision variables are sufficient to fully describe the optimal primal solution. Furthermore, the optimal dual variables can be computed using a stochastic approximation algorithm that is trivial to code, runs efficiently, and does not require a closed form for the cost function. Together, all of these developments answer an open question dating back to Dvoretzky et al. (1951), which proved the existence of fair partitions but did not offer any guidance as to how they may be computed.

Our case study sheds light on some of the practical challenges that may arise when considering the implementation of such a model. Fair representation can have unintended consequences, creating non-contiguous districts and significantly increasing travel times for the entire population, particularly for the very subgroups that policymakers aim to help. Nonetheless, our analysis can also be used to explore ways to mitigate these problems, for example by changing the cost function in a manner that reduces very high travel times. There are also difficult political and ethical considerations involved in determining the precise set of subgroups to include in the model. However, we believe that the main value of our paper is in offering a rigorous, evidence-based, politically neutral framework for assessing the feasibility of fair representation in complex regions that are often highly segregated to begin with.

References

Agueh, M. & Carlier, G. (2011), 'Barycenters in the Wasserstein space', *SIAM Journal on Mathematical Analysis* **43**(2), 904–924.

Aronov, B., Carmi, P. & Katz, M. J. (2009), 'Minimum-cost load-balancing partitions', *Algorithmica* **54**(3), 318–336.

Arredondo, V., Martínez-Panero, M., Pena, T. & Ricca, F. (2021), 'Mathematical political districting taking care of minority groups', *Annals of Operations Research* **305**(1), 375–402.

Aurenhammer, F. (1991), 'Voronoi diagrams – a survey of a fundamental geometric data structure', *ACM Computing Surveys* **23**(3), 345–405.

Aurenhammer, F. & Klein, R. (2000), Voronoi diagrams, *in* J.-R. Sack & J. Urrutia, eds, 'Handbook of Computational Geometry', Elsevier Science B. V., pp. 201–290.

Bach, F. & Moulines, E. (2011), Non-asymptotic analysis of stochastic approximation algorithms for machine learning, *in* J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira & K. Q. Weinberger, eds, 'Advances in Neural Information Processing Systems', Vol. 24, pp. 451–459.

Baker, B., Farrie, D., Johnson, M., Luhm, T. & Sciarra, D. G. (2017), 'Is school funding fair? a national report card', New Brunswick, NJ: Education Law Center.

Barocas, S. & Selbst, A. D. (2016), 'Big data's disparate impact', *California Law Review* **104**(3), 671–732.

Bayer, P., McMillan, R. & Rueben, K. S. (2004), 'What drives racial segregation? New evidence using census microdata', *Journal of Urban Economics* **56**(3), 514–535.

Brill, S. (2019), 'The law of school catchment areas', *Stanford Law & Policy Review* **30**(2), 349–398.

Calders, T. & Verwer, S. (2010), 'Three naive Bayes approaches for discrimination-free classification', *Data Mining and Knowledge Discovery* **21**(2), 277–292.

Carlsson, J. G., Carlsson, E. & Devulapalli, R. (2016), 'Shadow prices in territory division', *Networks and Spatial Economics* **16**(3), 893–931.

Carlsson, J. G. & Devulapalli, R. (2013), 'Dividing a territory among several facilities', *INFORMS Journal on Computing* **25**(4), 730–742.

Carlsson, J. G., Peng, X. & Ryzhov, I. O. (2024), 'Demand equilibria in spatial service systems', *Manufacturing & Service Operations Management* **26**(6), 2305–2321.

Castro, A., Siegel-Hawley, G., Bridges, K. & Williams, S. E. (2024), 'Drawn into policy: A systematic review of school rezoning rationales, processes, and outcomes', *Review of Educational Research* **94**(4), 539–583.

Chouldechova, A. (2017), 'Fair prediction with disparate impact: A study of bias in recidivism prediction instruments', *Big Data* **5**(2), 153–163.

Chzhen, E., Denis, C., Hebiri, M., Oneto, L. & Pontil, M. (2020), Fair regression with Wasserstein barycenters, *in* H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan & H. Lin, eds, 'Advances in Neural Information Processing Systems', Vol. 33, pp. 7321–7331.

Cinlar, E. (2011), *Probability and Stochastics*, Springer.

Cohen-Addad, V., Klein, P. N. & Young, N. E. (2018), Balanced centroidal power diagrams for redistricting, *in* F. Banaei-Kashani, E. Hoel, R. H. Güting, R. Tamassia & L. Xiong, eds, 'Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems', pp. 389–396.

del Barrio, E., Gordaliza, P. & Loubes, J.-M. (2020), 'Review of mathematical frameworks for fairness in machine learning', *arXiv preprint arXiv:2005.13755* .

Dvoretzky, A., Wald, A. & Wolfowitz, J. (1951), 'Relations among certain ranges of vector measures', *Pacific Journal of Mathematics* **1**(1), 59–74.

Dwork, C., Hardt, M., Pitassi, T., Reingold, O. & Zemel, R. (2012), Fairness through awareness, *in* S. Goldwasser, ed., 'Proceedings of the 3rd Conference on Innovations in Theoretical Computer Science', pp. 214–226.

ESRI (2025), 'ArcGIS Online', https://www.arcgis.com/home/item.html?id=414f777693de44dbb0667eb001dc1 San Fernando Valley boundary shapefile.

Genevay, A., Cuturi, M., Peyré, G. & Bach, F. (2016), Stochastic optimization for large-scale optimal transport, *in* D. Lee, M. Sugiyama, U. von Luxburg, I. Guyon & R. Garnett, eds, 'Advances in Neural Information Processing Systems', Vol. 29, Curran Associates, Inc., pp. 3440–3448.

Gomilla, K. & Arangala, C. (2020), 'Mathematical modeling for Charlotte-Mecklenburg school zones', *The North Carolina Journal of Mathematics and Statistics* **6**(1), 14–30.

Gupta, S. & Kamble, V. (2021), 'Individual fairness in hindsight', *Journal of Machine Learning Research* **22**, 144:1–144:35.

Hardt, M., Price, E., Price, E. & Srebro, N. (2016), Equality of opportunity in supervised learning, *in* D. Lee, M. Sugiyama, U. von Luxburg, I. Guyon & R. Garnett, eds, 'Advances in Neural Information Processing Systems', Vol. 29, Curran Associates, Inc., pp. 3315–3323.

Hartmann, V. & Schuhmacher, D. (2020), 'Semi-discrete optimal transport: a solution procedure for the unsquared Euclidean distance case', *Mathematical Methods of Operations Research* **92**(1), 133–163.

Haugland, D., Ho, S. C. & Laporte, G. (2007), 'Designing delivery districts for the vehicle routing problem with stochastic demands', *European Journal of Operational Research* **180**(3), 997–1010.

Johndrow, J. E. & Lum, K. (2019), 'An algorithm for removing sensitive information: application to race-independent recidivism prediction', *The Annals of Applied Statistics* **13**(1), 189–220.

Joseph, M., Kearns, M., Morgenstern, J. H. & Roth, A. (2016), Fairness in learning: Classic and contextual bandits, *in* D. Lee, M. Sugiyama, U. von Luxburg, I. Guyon & R. Garnett, eds, 'Advances in Neural Information Processing Systems', Vol. 29, pp. 325–333.

Kallus, N. & Zhou, A. (2021), Fairness, welfare, and equity in personalized pricing, *in* 'Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency', pp. 296–314.

Kleinberg, J., Mullainathan, S. & Raghavan, M. (2017), Inherent trade-offs in the fair determination of risk scores, *in* C. H. Papadimitriou, ed., 'Proceedings of the 8th Conference on Innovations in Theoretical Computer Science', pp. 43:1–43:23.

Kushner, H. J. & Yin, G. (2003), *Stochastic approximation and recursive algorithms and applications (2nd ed.)*, Springer.

Luenberger, D. G. (1997), *Optimization by vector space methods*, John Wiley & Sons.

NCES (2022), 'National Center for Education Statistics, Education Demographic and Geographic Estimates Program', https://nces.ed.gov/programs/edge/geographic/schoollocations. Local Education Agency geocoded data files.

Owens, A. (2016), 'Inequality in children's contexts: Income segregation of households with and without children', *American Sociological Review* **81**(3), 549–574.

Owens, A. (2018), 'Income segregation between school districts and inequality in students' achievement', *Sociology of Education* **91**(1), 1–27.

Owens, A., Reardon, S. F. & Jencks, C. (2016), 'Income segregation between schools and school districts', *American Educational Research Journal* **53**(4), 1159–1197.

Ozel, A., Smilowitz, K. & Goldstein, L. K. S. (2025), 'Community-engaged school district design: a stream-based approach', *Operations Research* **73**(4), 1916–1931.

Pang, V. O., Han, P. P. & Pang, J. M. (2011), 'Asian American and Pacific Islander students: Equity and the achievement gap', *Educational Researcher* **40**(8), 378–389.

Pasupathy, R. & Kim, S. (2011), 'The stochastic root-finding problem: Overview, solutions, and open questions', *ACM Transactions on Modeling and Computer Simulation* **21**(3), 19:1–19:23.

Pearce, J. (2000), 'Techniques for defining school catchment areas for comparison with census data', *Computers, Environment and Urban Systems* **24**(4), 283–303.

Peyré, G. & Cuturi, M. (2019), 'Computational optimal transport: With applications to data science', *Foundations and Trends in Machine Learning* **11**(5-6), 355–607.

Polyak, B. T. & Juditsky, A. B. (1992), 'Acceleration of stochastic approximation by averaging', *SIAM Journal on Control and Optimization* **30**(4), 838–855.

Pratelli, A. (2007), 'On the equality between Monge's infimum and Kantorovich's minimum in optimal mass transportation', *Annales de l'Institut Henri Poincare* **B43**(1), 1–13.

Reardon, S. F. & Bischoff, K. (2011), 'Income inequality and income segregation', *American Journal of Sociology* **116**(4), 1092–1153.

Ricca, F., Scozzari, A. & Simeone, B. (2008), 'Weighted Voronoi region algorithms for political districting', *Mathematical and Computer Modelling* **48**(9-10), 1468–1477.

Ryabov, I. & Van Hook, J. (2007), 'School segregation and academic achievement among Hispanic children', *Social Science Research* **36**(2), 767–788.

Saporito, S. (2017*a*), 'Irregularly-shaped school attendance zones and racial integration', *Social Science Research* **64**, 299–318.

Saporito, S. (2017*b*), 'Shaping income segregation in schools: The role of school attendance zone geography', *American Educational Research Journal* **54**(6), 1345–1377.

Shahmizad, M. & Buchanan, A. (2025), 'Political districting to minimize county splits', *Operations Research* **73**(2), 752–774.

Swamy, R., King, D. M. & Jacobson, S. H. (2023), 'Multiobjective optimization for politically fair districting: A scalable multilevel approach', *Operations Research* **71**(2), 536–562.

U.S. Supreme Court (2007), 'Parents Involved in Community Schools v. Seattle School District No. 1', 551 US 701.

U.S. Supreme Court (2023), 'Students for Fair Admissions, Inc. v. President and Fellows of Harvard College', 600 US 181.

Villani, C. (2021), *Topics in optimal transportation*, American Mathematical Society.

Yushimito, W. F., Jaller, M. & Ukkusuri, S. (2012), 'A Voronoi-based heuristic algorithm for locating distribution centers in disasters', *Networks and Spatial Economics* **12**(1), 21–39.

Zhu, Y. & Ryzhov, I. O. (2025*a*), 'Optimal data-driven hiring with equity for underrepresented groups', *Production and Operations Management (to appear)* .

Zhu, Y. & Ryzhov, I. O. (2025*b*), 'Semidiscrete optimal transport with unknown costs', *arXiv preprint arXiv:2310.00786* .

Zink, A. & Rose, S. (2020), 'Fair regression for health care spending', *Biometrics* **76**(3), 973–982.

8 Appendix: proofs

Below, we provide complete proofs for all results that were stated in the text.

8.1 Proof of Theorem 4.1

The proof uses the following result from Sec. 8.6 of Luenberger (1997).

Theorem 8.1. *Let $\mathfrak{f}$ be a real-valued convex functional defined on a convex subset Ω of a vector space $\mathfrak{X}$, and let $\mathfrak{G}$ be a convex mapping of $\mathfrak{X}$ into a normed space $\mathfrak{Z}$. Suppose there exists an $\mathfrak{x}_1$ such that $\mathfrak{G}(\mathfrak{x}_1) < \theta$, where θ is the zero element, and that $\inf\{\mathfrak{f}(x) : \mathfrak{x} \in \Omega, \mathfrak{G}(\mathfrak{x}) \leq \theta\}$ is finite. Then*

$$\inf_{\mathfrak{x} \in \Omega, \, \mathfrak{G}(\mathfrak{x}) \leq \theta} \mathfrak{f}(\mathfrak{x}) \quad = \quad \max_{\mathfrak{z}^* \geq \theta} \inf_{\mathfrak{x} \in \Omega} \mathfrak{f}(\mathfrak{x}) + \langle \mathfrak{G}(\mathfrak{x}), \mathfrak{z}^* \rangle$$

and the maximum on the right is achieved by some $\mathfrak{z}_0^* \geq \theta$. If the infimum on the left is achieved by some $\mathfrak{x}_0 \in \Omega$, then $\langle \mathfrak{G}(\mathfrak{x}_0), \mathfrak{z}_0^* \rangle = 0$, and $\mathfrak{x}_0$ minimizes $\mathfrak{f}(\mathfrak{x}) + \langle \mathfrak{G}(\mathfrak{x}), \mathfrak{z}_0^* \rangle$ over all $\mathfrak{x} \in \Omega$.

Consider the problem

$$\max_{u_{kz}} \iint_{\mathcal{X}} \sum_z f_z(x) \min_k \{q_z c(x,k) - f_z(x) u_{kz}\} + \sum_{k,z} p_k u_{kz} \tag{30}$$

which has a maximizer u_{kz}^* by compactness of $\mathcal{X}$. We may rewrite (30) as an LP over $\mathbb{R}^{KM} \oplus L_1(\mathcal{X})$. This problem has the objective

$$\max_{u_{kz}, \sigma(x)} \iint_{\mathcal{X}} \sigma(x) + \sum_{k,z} p_k u_{kz}$$

subject to the constraints

$$\sum_z f_z(x) u_{kz} + \sigma(x) \leq \left(\sum_z q_z f_z(x) \right) c(x,k) \qquad \forall x, k.$$

Equivalently, we may state it as

$$\min_{u_{kz}, \sigma(x)} - \left(\iint_{\mathcal{X}} \sigma(x) + \sum_{k,z} p_k u_{kz} \right)$$

subject to

$$\sum_z f_z(x) u_{kz} + \sigma(x) - \left(\sum_z q_z f_z(x) \right) c(x,k) \leq 0 \qquad \forall x, k.$$

We thus have

$$\mathfrak{X} = \Omega = \mathbb{R}^{KZ} \oplus L_1(\mathcal{X})$$

$$\mathfrak{x} = \begin{pmatrix} u_{kz} \\ \sigma(\cdot) \end{pmatrix}$$

$$\mathfrak{Z} = L^1(\mathcal{X})^K$$

$$\mathfrak{G}(\mathfrak{x}) = \sum_z f_z(x) u_{kz} + \sigma(x) - \left(\sum_z q_z f_z(x) \right) c(x, k)$$

According to Theorem 8.1, the solution exists and we have $\mathfrak{Z}^* = L^\infty(\mathcal{X})^K$, so $\mathfrak{z}^* = g(x, k)$ and the solution is the same as

$$\max_{g(\cdot,\cdot) \geq 0} \inf_{u_{kz}, \sigma(x)} \quad - \left(\iint_{\mathcal{X}} \sigma(x) \, dx + \sum_{k,z} p_k u_{kz} \right)$$
$$+ \sum_k \iint_{\mathcal{X}} \left[\left(\sum_z f_z(x) u_{kz} \right) + \sigma(x) - \left(\sum_z q_z f_z(x) \right) c(x, k) \right] g(x, k) \, dx.$$

Equivalently, we may write

$$\max_{g(\cdot,\cdot) \geq 0} \inf_{u_{kz}, \sigma(x)} - \sum_{k,z} p_k u_{kz} + \iint_{\mathcal{X}} \sum_k \left[\left(\sum_z f_z(x) u_{kz} \right) + \sigma(x) - \left(\sum_z q_z f_z(x) \right) c(x, k) \right] g(x, k) - \sigma(x) \, dx,$$

which is the same as

$$\max_{g(\cdot,\cdot) \geq 0} \inf_{u_{kz}, \sigma(x)} \iint_{\mathcal{X}} \sum_k \left[- \left(\sum_z q_z f_z(x) \right) c(x, k) \right] g(x, k) \, dx$$
$$+ \iint_{\mathcal{X}} \left(\sum_k g(x, k) - 1 \right) \sigma(x) \, dx + \iint_{\mathcal{X}} \sum_{k,z} \left(f_z(x) g(x, k) - p_k \right) u_{kz} \, dx. \tag{31}$$

Because we are taking $\inf_{u_{kz}, \sigma(x)}$, (31) requires $\sum_k g(x, k) = 1 \, \forall x \in \mathcal{X}$, and $\iint_{\mathcal{X}} g(x, k) f_z(x) \, dx = p_k \, \forall k, \forall z$. Thus, swapping out the minus sign in the objective function and replacing it with a minimization, the problem is the same as

$$\min_g \sum_k \iint_{\mathcal{X}} \sum_z q_z c(x, k) g(x, k) f_z(x) \, dx$$

subject to

$$\sum_k g(x, k) = 1 \qquad \forall x \in \mathcal{X}$$

$$\iint_{\mathcal{X}} g(x, k) f_z(x) \, dx = p_k \qquad \forall k, \forall z$$

$$g \geq 0$$

as desired.

8.2   Proof of Theorem 5.1

Observing that $v^{n+1} - v^* = v^n - v^* + \alpha_n \zeta^n (v^n)$, we write

$$
\begin{aligned}
\mathbb{E}\left(\|v^{n+1} - v^*\|_2^2\right) &= \mathbb{E}\left(\|v^n - v^*\|_2^2\right) + \alpha_n^2 \mathbb{E}\left(\|\zeta^n(v^n)\|_2^2\right) + 2\alpha_n \mathbb{E}\left((v^n - v^*)^\top \nabla_v h(v^n)\right) \quad (32) \\
&\leq \mathbb{E}\left(\|v^n - v^*\|_2^2\right) + \alpha_n^2 C^2, \quad (33)
\end{aligned}
$$

where (32) is obtained by conditioning the last term on v^n, and (33) follows by the boundedness of ζ^n and the concavity of h, i.e., $0 \leq h(v^*) - h(v^n) \leq \nabla_v h(v^n)^\top (v^* - v^n)$. Thus, we have

$$
\mathbb{E}\left(\|v^n - v^*\|_2^2\right) \leq \|v^0 - v^*\|_2^2 + C^2 \sum_{m=0}^{n-1} \alpha_m^2 \leq C_1 \log n
$$

for some sufficiently large $C_1 > 0$.

Applying the concavity of h twice, we derive

$$
\begin{aligned}
h(v^*) - h\left(\frac{1}{n}\sum_{m=1}^{n} v^m\right) &\leq h(v^*) - \frac{1}{n}\sum_{m=1}^{n} h(v^m) \\
&= \frac{1}{n}\sum_{m=1}^{n} h(v^*) - h(v^m) \\
&\leq \frac{1}{n}\sum_{m=1}^{n} (v^* - v^m)^\top \nabla_v h(v^m). \quad (34)
\end{aligned}
$$

By (32), we have

$$
\mathbb{E}\left((v^* - v^m)^\top \nabla_v h(v^m)\right) = \frac{\delta^m - \delta^{m+1} + \alpha_m^2 \mathbb{E}\left(\|\zeta^m(v^m)\|_2^2\right)}{2\alpha_m}, \quad (35)
$$

where $\delta^m = \mathbb{E}\left(\|v^m - v^*\|_2^2\right)$. Plugging (35) into (34) yields

$$
\begin{aligned}
h(v^*) - \mathbb{E}\left(h\left(\frac{1}{n}\sum_{m=1}^{n} v^m\right)\right) &\leq \frac{1}{n}\sum_{m=1}^{n} \frac{\delta^m - \delta^{m+1} + \alpha_m^2 \mathbb{E}\left(\|\zeta^m(v^m)\|_2^2\right)}{2\alpha_m} \\
&\leq \frac{1}{2n}\sum_{m=1}^{n} \frac{\delta^m - \delta^{m+1}}{\alpha_m} + \frac{C^2}{2n}\sum_{m=1}^{n} \alpha_m \\
&= \frac{1}{2n}\left(\sum_{m=1}^{n} \frac{\delta^m}{\alpha_m} - \sum_{m=1}^{n} \frac{\delta^{m+1}}{\alpha_m} + \sum_{m=1}^{n} \delta^{m+1}\left(\frac{1}{\alpha_{m+1}} - \frac{1}{\alpha_m}\right)\right) + \frac{C^2}{2n}\sum_{m=1}^{n} \alpha_m \\
&= \frac{1}{2n}\left(\frac{\delta^1}{\alpha_1} - \frac{\delta^{n+1}}{\alpha_{n+1}} + \sum_{m=1}^{n} \delta^{m+1}\left(\frac{1}{\alpha_{m+1}} - \frac{1}{\alpha_m}\right)\right) + \frac{C^2}{2n}\sum_{m=1}^{n} \alpha_m \\
&\leq \frac{1}{2n}\left(\frac{\delta^1}{\alpha_1} - \frac{\delta^{n+1}}{\alpha_{n+1}} + C_1 \sum_{m=1}^{n}\left(\frac{1}{\alpha_{m+1}} - \frac{1}{\alpha_m}\right)\log m\right) + \frac{C^2}{2n}\sum_{m=1}^{n} \alpha_m
\end{aligned}
$$

$$\leq \quad \frac{1}{2n}\left(\frac{\delta^1}{\alpha_1} - \frac{\delta^{n+1}}{\alpha_{n+1}} + C_1 \log n \sum_{m=1}^{n} \frac{1}{\alpha_{m+1}} - \frac{1}{\alpha_m}\right) + \frac{C^2}{2n}\sum_{m=1}^{n}\alpha_m$$

$$= \quad \frac{1}{2n}\left(\frac{\delta^1}{\alpha_1} - \frac{\delta^{n+1}}{\alpha_{n+1}} + C_1 \log n \left(\frac{1}{\alpha_{n+1}} - \frac{1}{\alpha_1}\right)\right) + \frac{C^2}{2n}\sum_{m=1}^{n}\alpha_m$$

$$\leq \quad \frac{C_1}{2n}\cdot\frac{1}{\alpha_{n+1}}\cdot\log n + \frac{C^2}{2n}\sum_{m=1}^{n}\alpha_m.$$

The desired result follows after plugging in the definition of $\{\alpha_n\}$.

9 Appendix: complete visuals for case study

In this section, we present a complete set of plots showing the district boundaries for all 78 facilities in Sections 6.1-6.2. The plots were generated in the same way as in Figures 4-6. A blank plot indicates that the district office in question was closed, i.e., assigned $p_k^* = 0$ by the relevant optimization model.

There are 312 plots in all, presented in groups of 20. We first provide plots for the seven-group setting from Section 6.1. The first set of 78 is for regular (unsquared) travel distance as the cost function, and the second set is for squared distance. We then provide two more sets for the three-group case (Section 6.2).

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

(k)

(l)

(m)

(n)

(o)

(p)

(q)

(r)

(s)

(t)
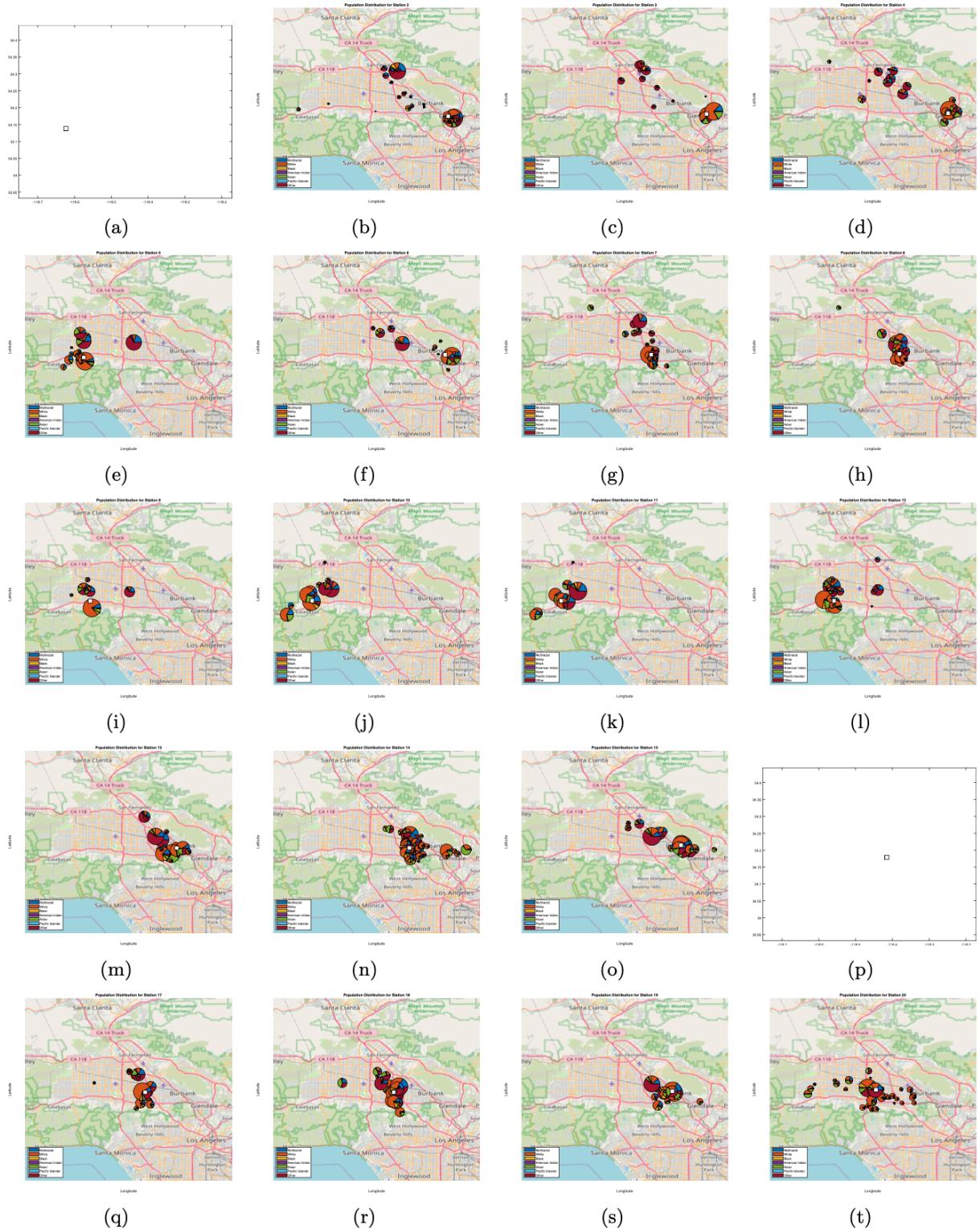
Figure 7: District boundaries for district offices 1-20 under regular distance (seven groups).
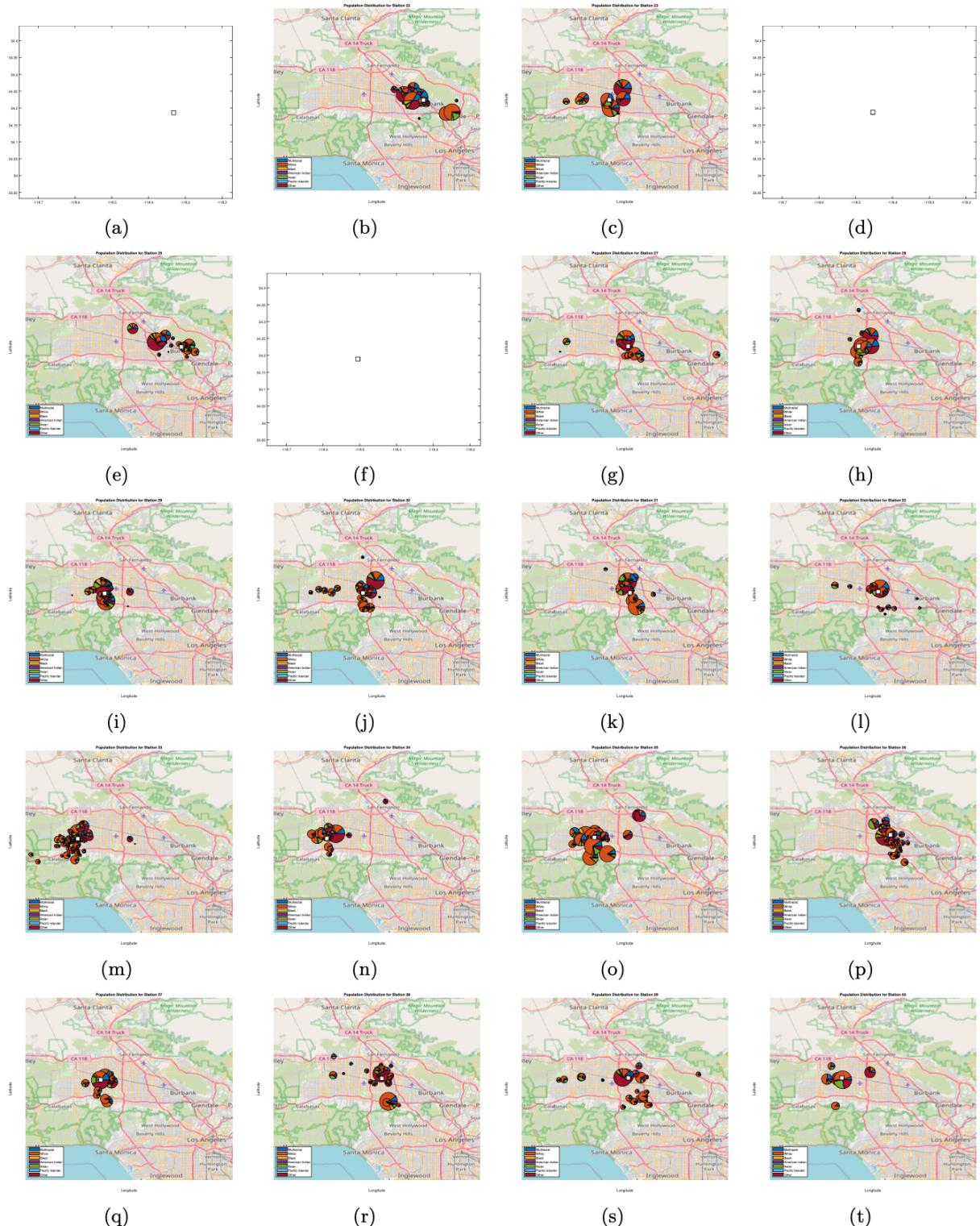
Figure 8: District boundaries for district offices 21-40 under regular distance (seven groups).
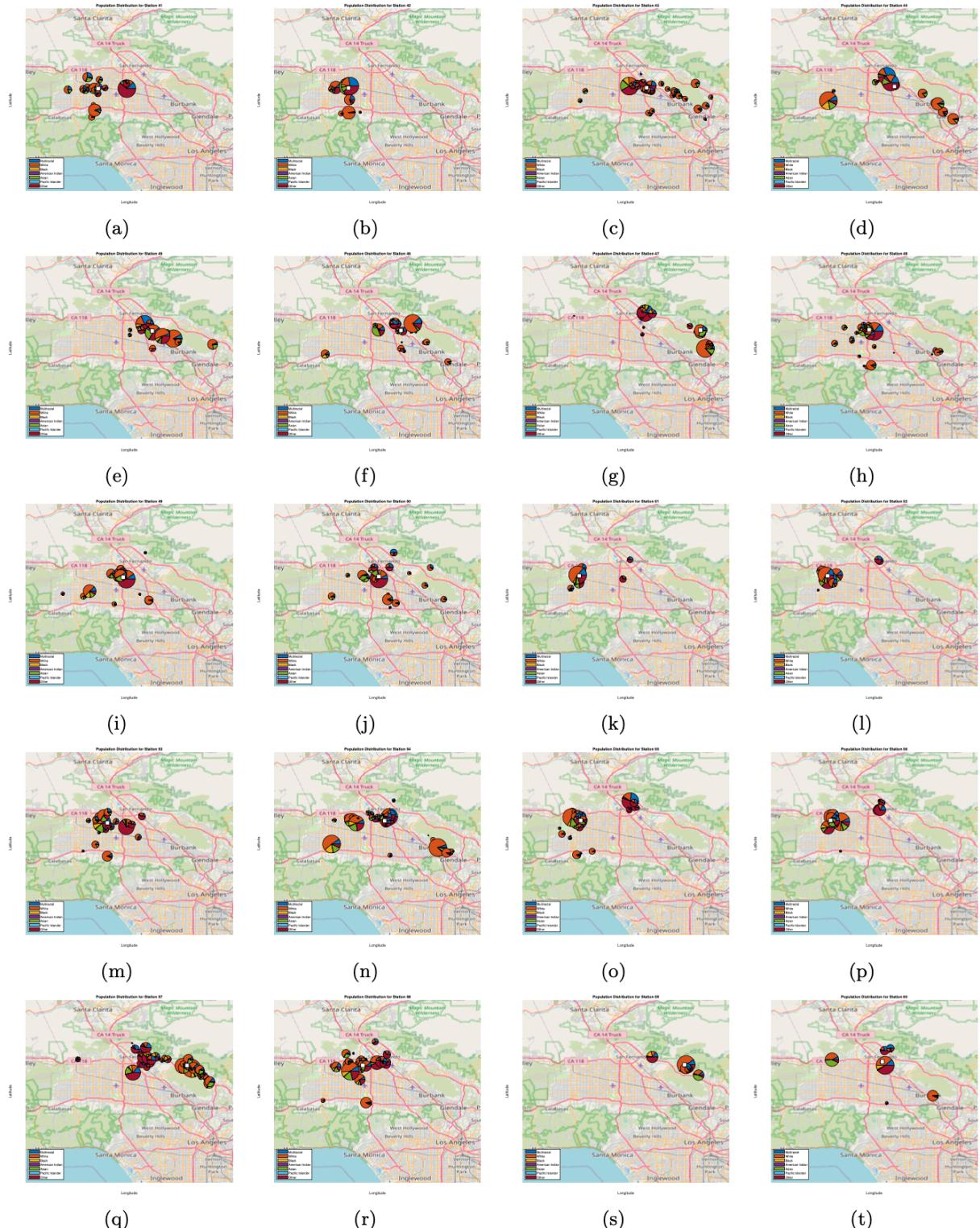
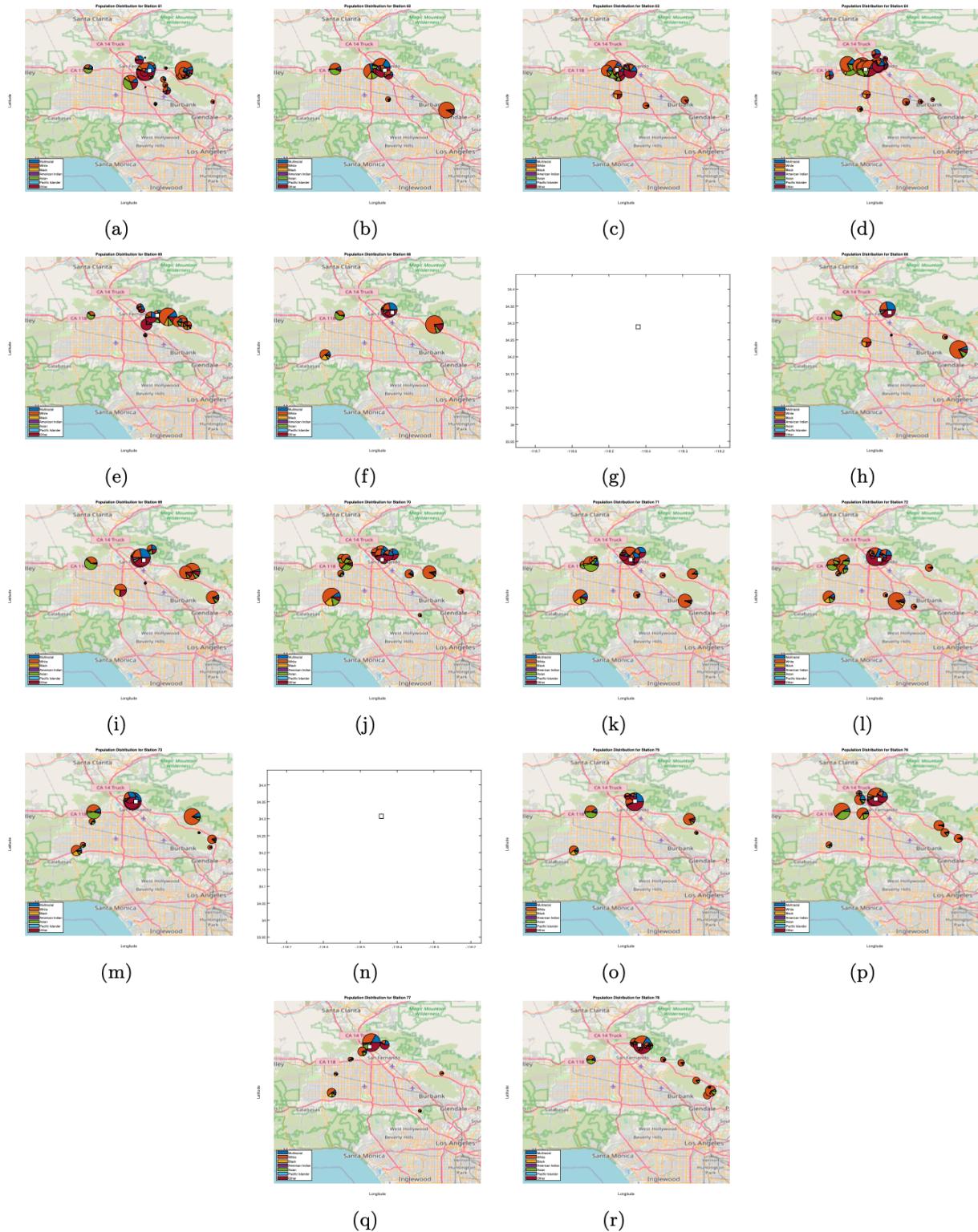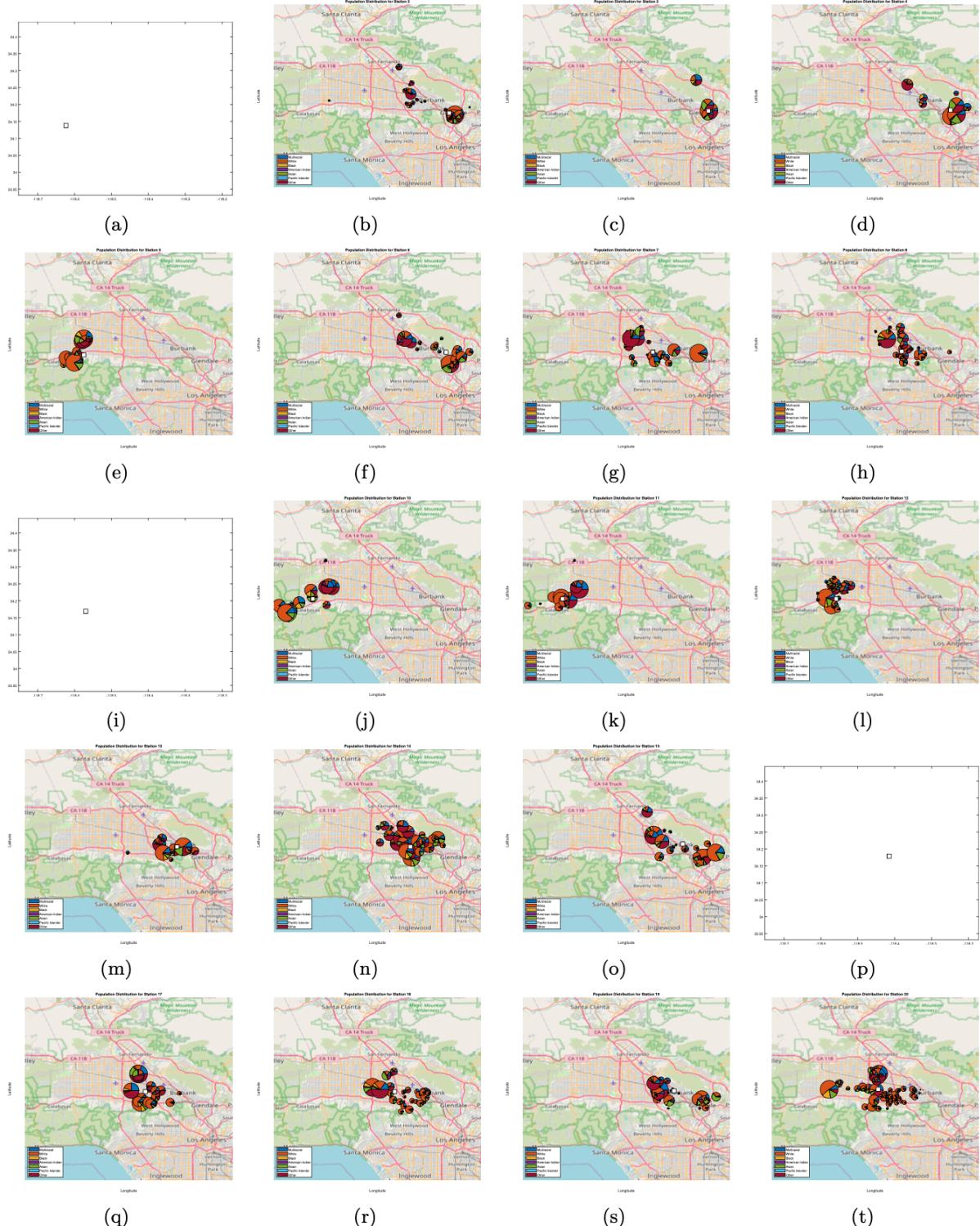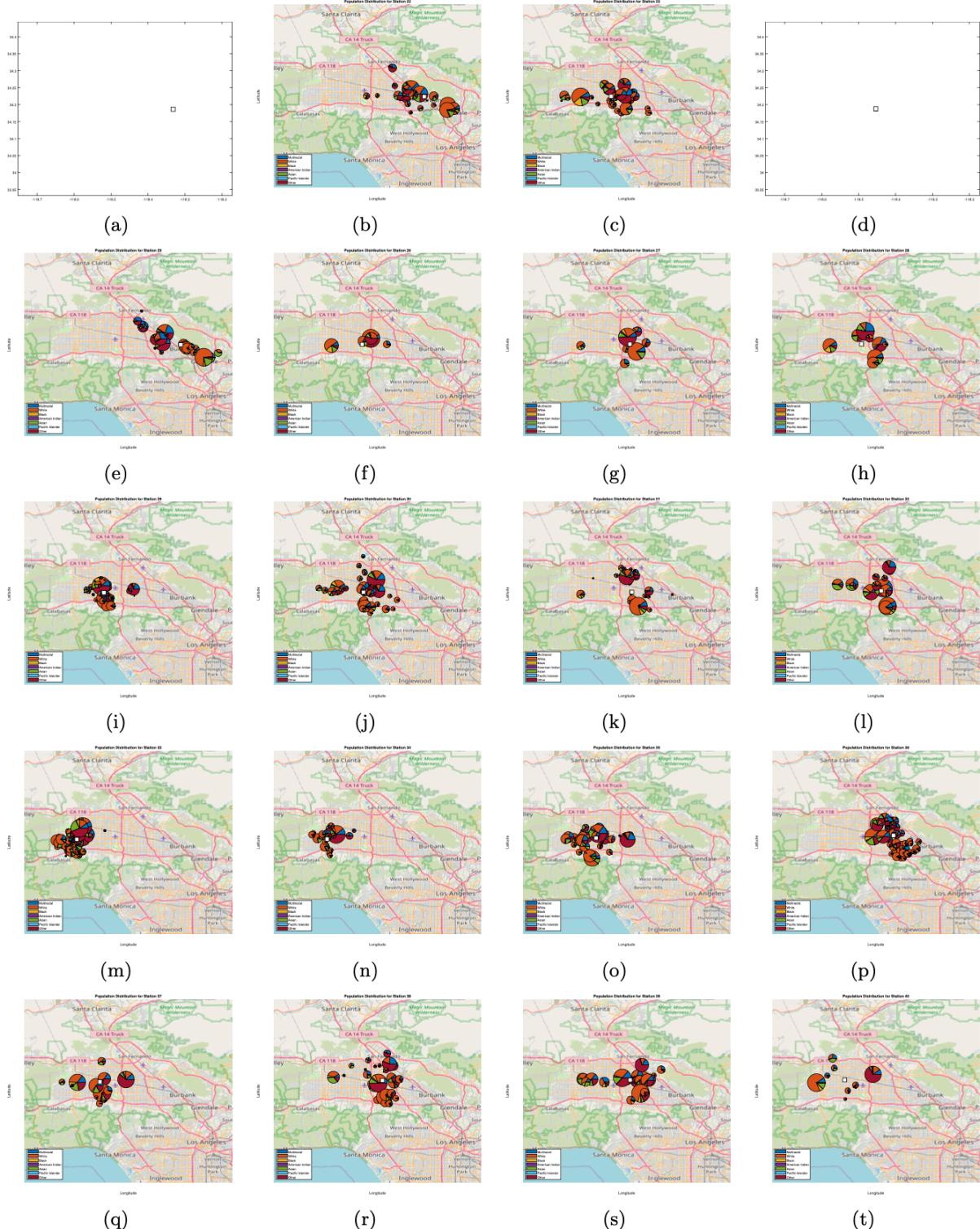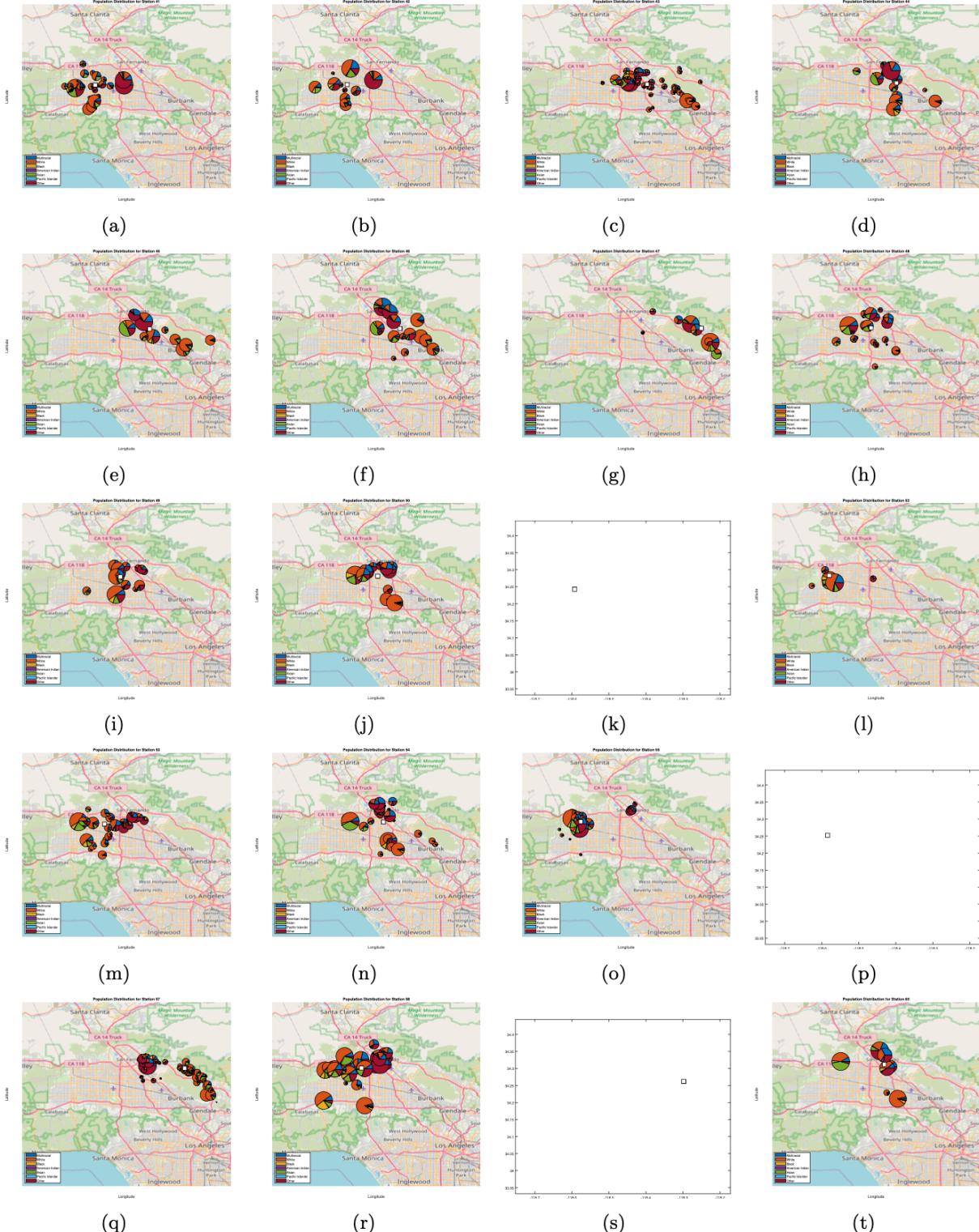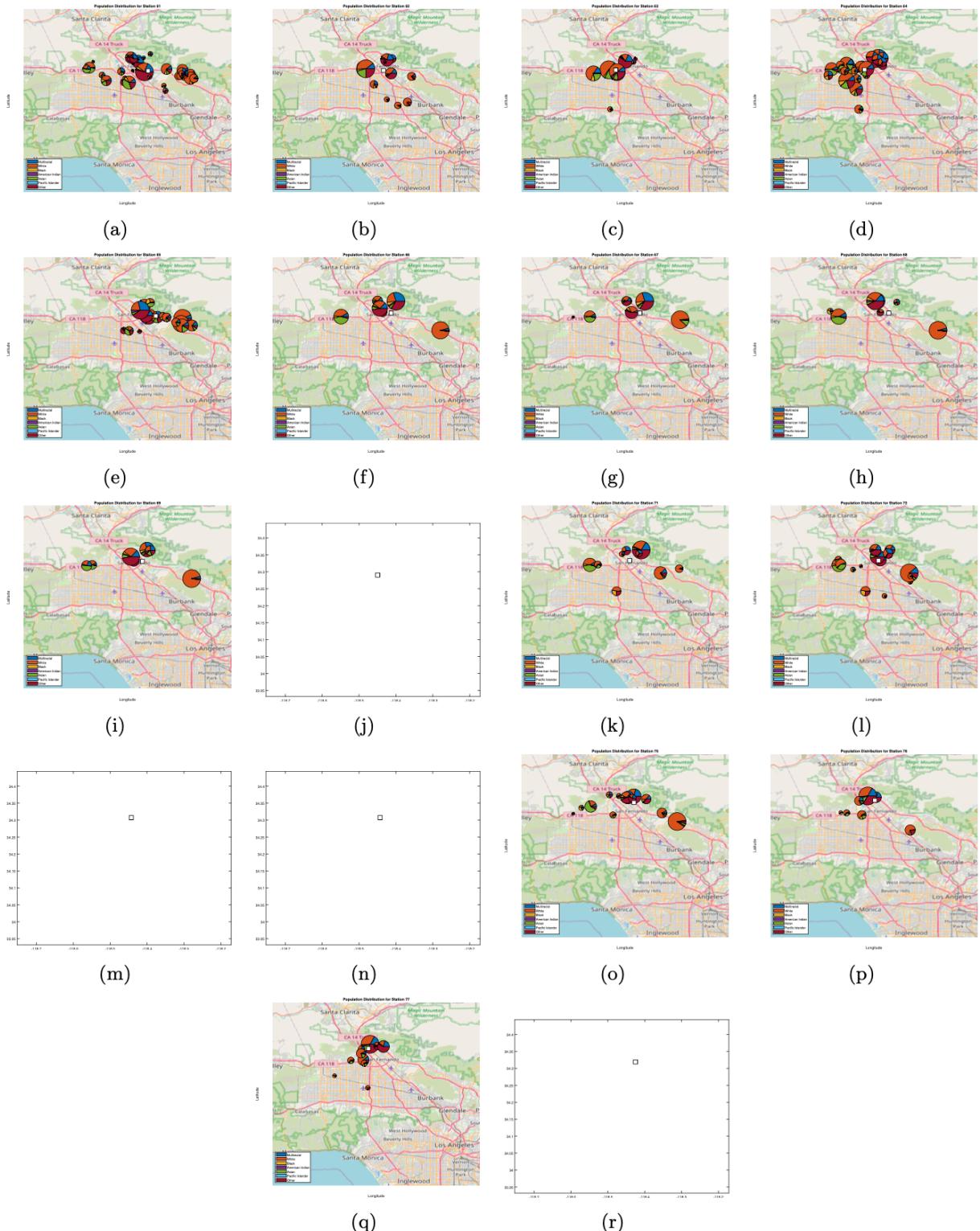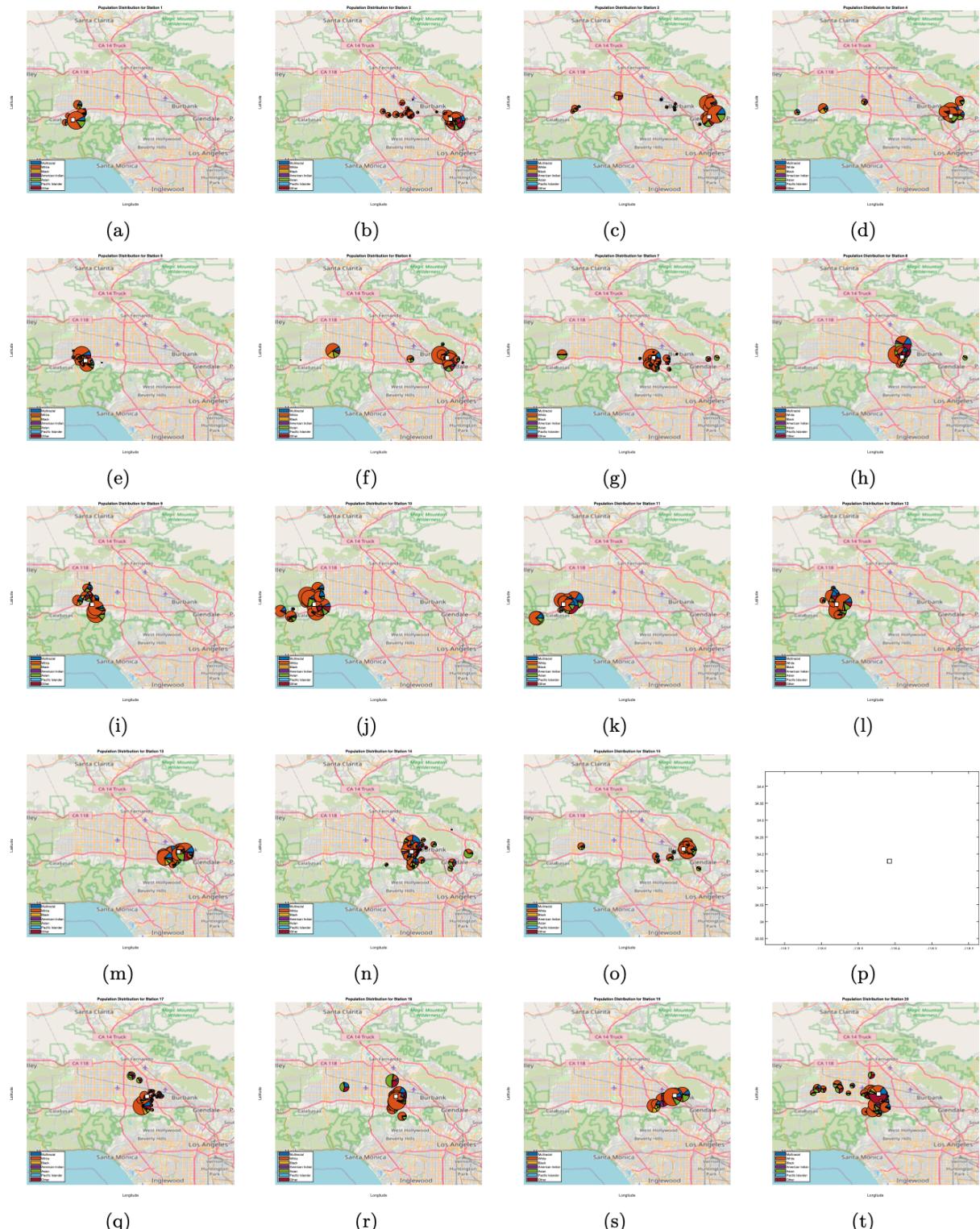Figure 9: District boundaries for district offices 41-60 under regular distance (seven groups).
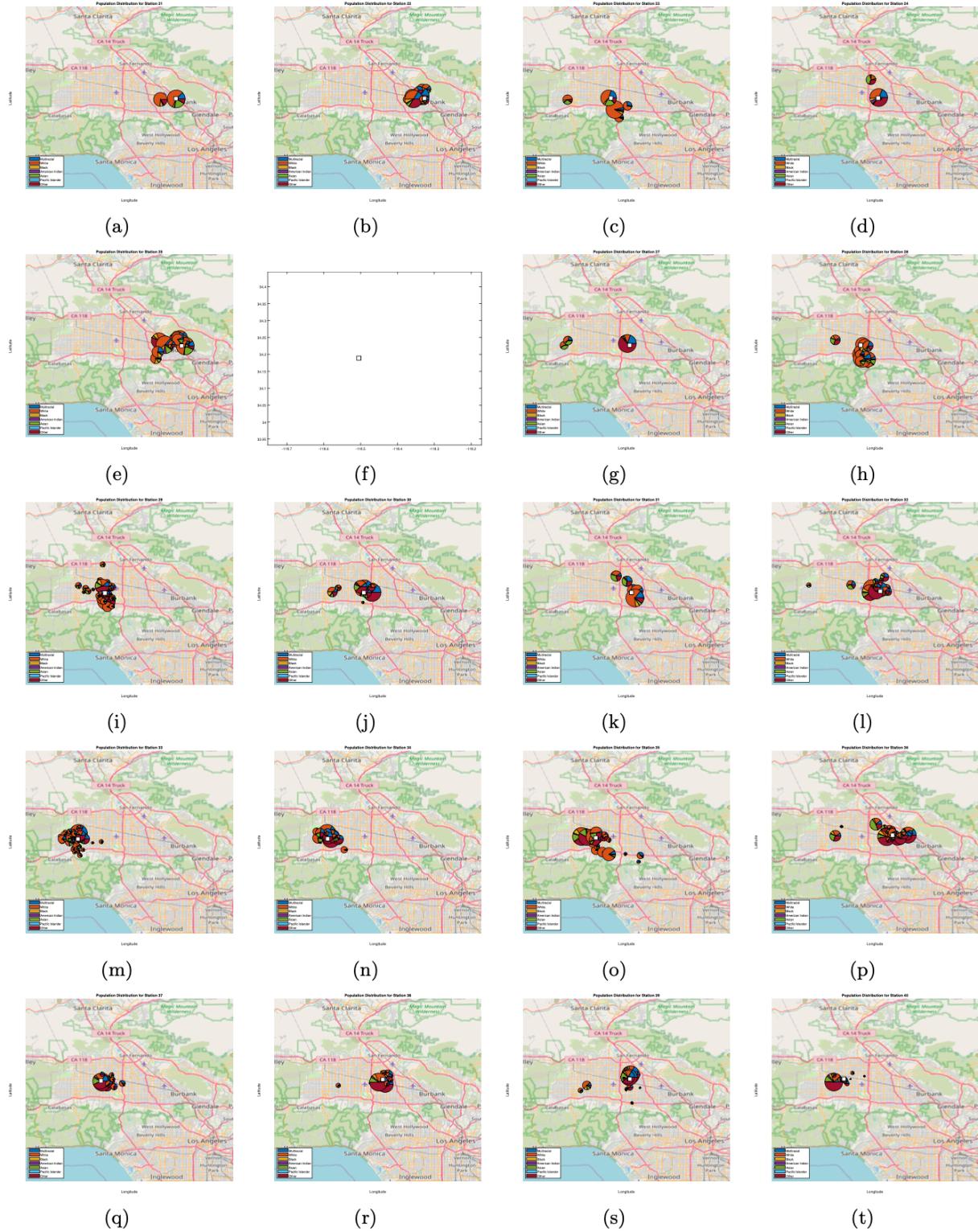
Figure 10: District boundaries for district offices 61-78 under regular distance (seven groups).

Figure 11: District boundaries for district offices 1-20 under squared distance (seven groups).

Figure 12: District boundaries for district offices 21-40 under squared distance (seven groups).
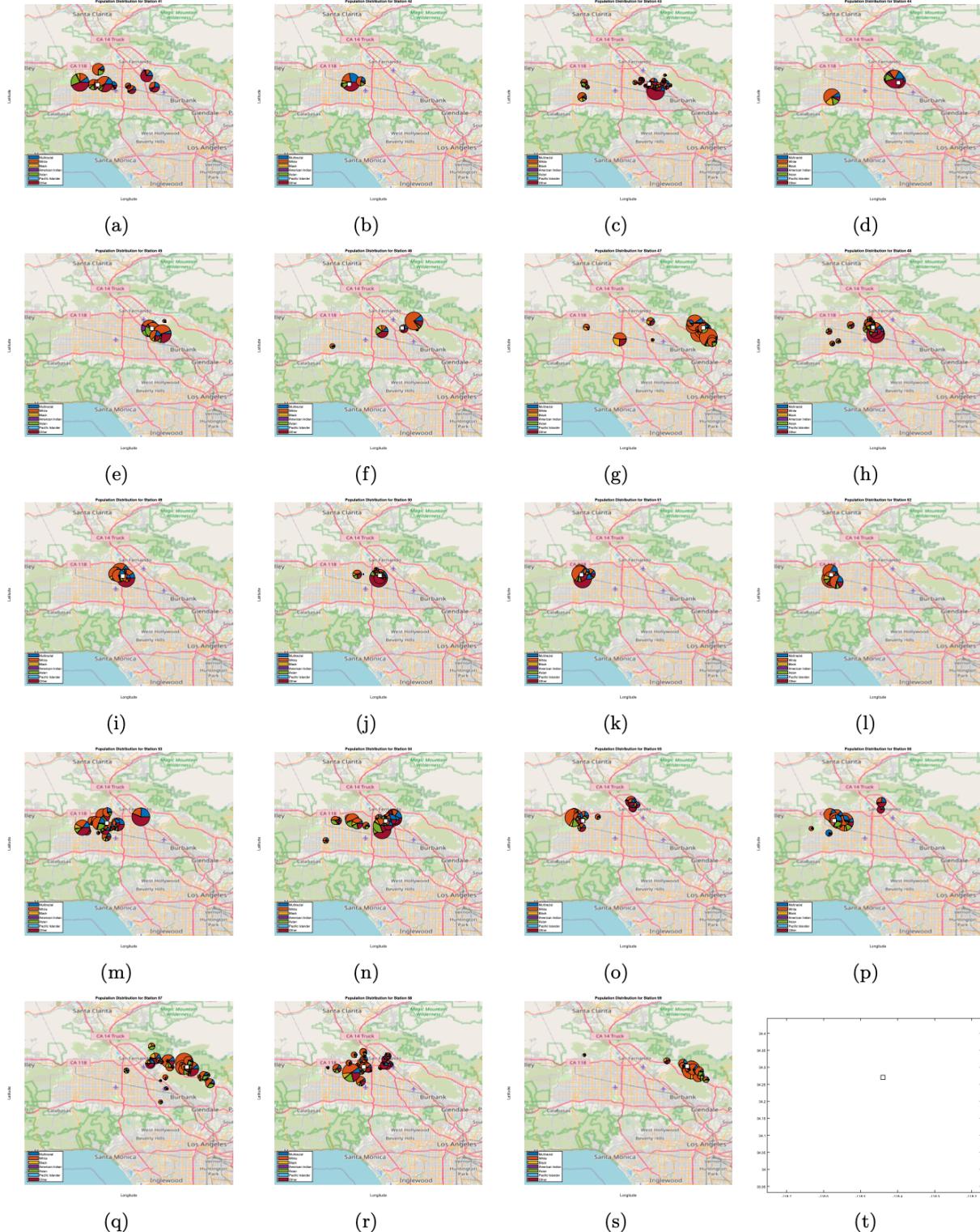
Figure 13: District boundaries for district offices 41-60 under squared distance (seven groups).

Figure 14: District boundaries for district offices 61-78 under squared distance (seven groups).

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

(k)

(l)

(m)

(n)

(o)

(p)

(q)

(r)

(s)

(t)

Figure 15: District boundaries for district offices 1-20 under regular distance (three groups).

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

(e) (f) (g) (h)

(i) (j) (k) (l)

(m) (n) (o) (p)

(q) (r) (s) (t)

Figure 16: District boundaries for district offices 21-40 under regular distance (three groups).

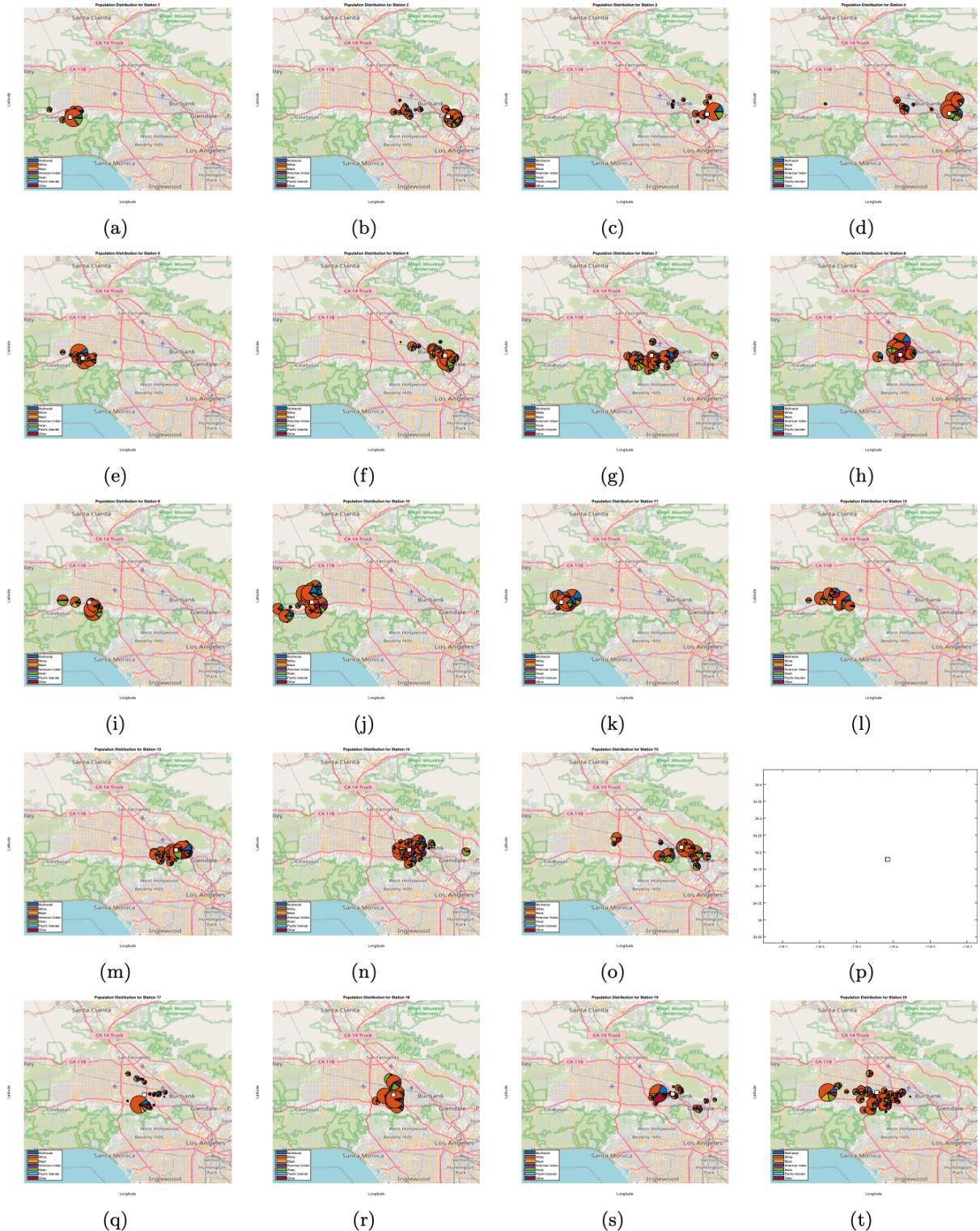Figure 17: District boundaries for district offices 41-60 under regular distance (three groups).

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

(k)

(l)

(m)

(n)

(o)

(p)

(q)

(r)

Figure 18: District boundaries for district offices 61-78 under regular distance (three groups).

Figure 19: District boundaries for district offices 1-20 under squared distance (three groups).
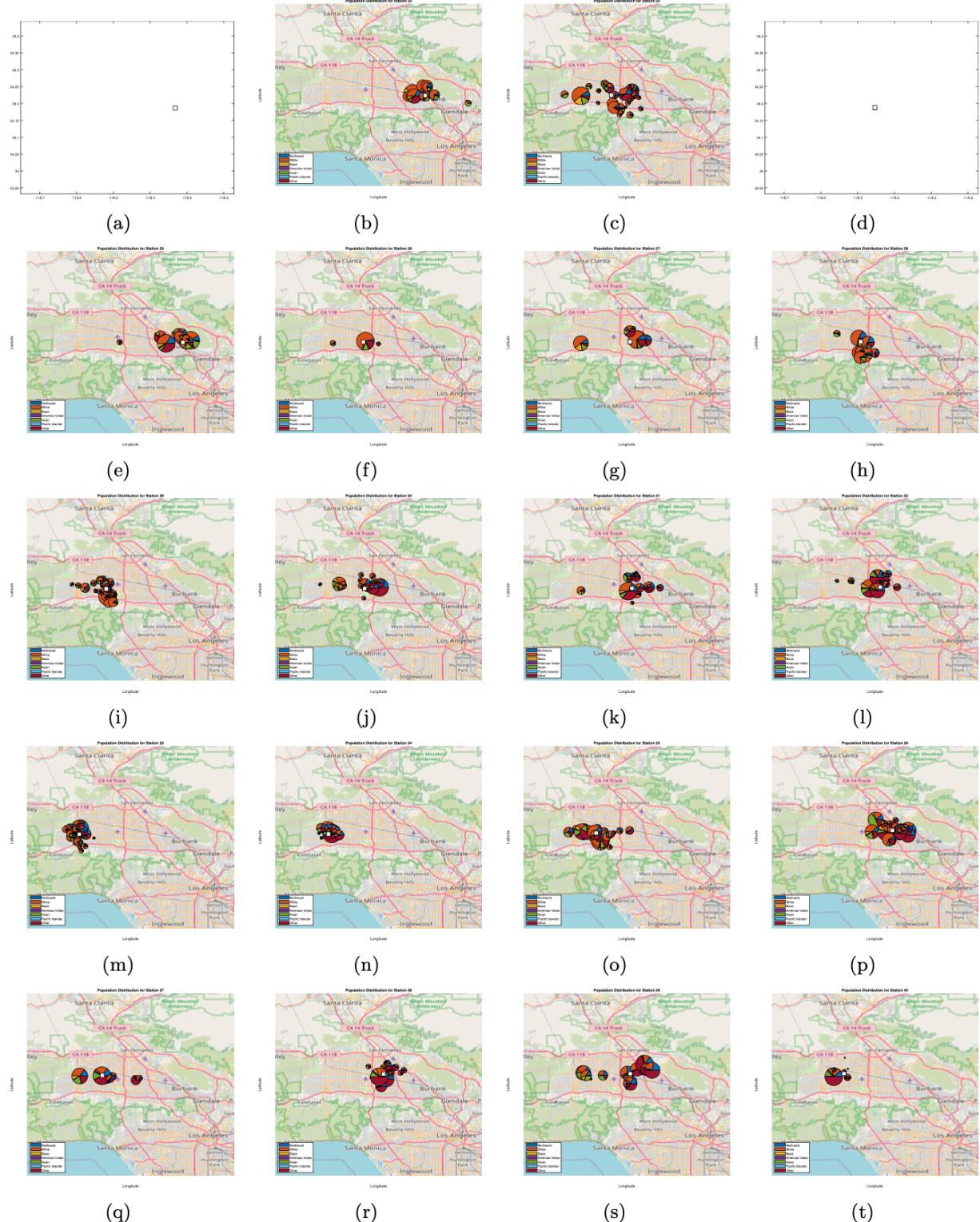
Figure 20: District boundaries for district offices 21-40 under squared distance (three groups).
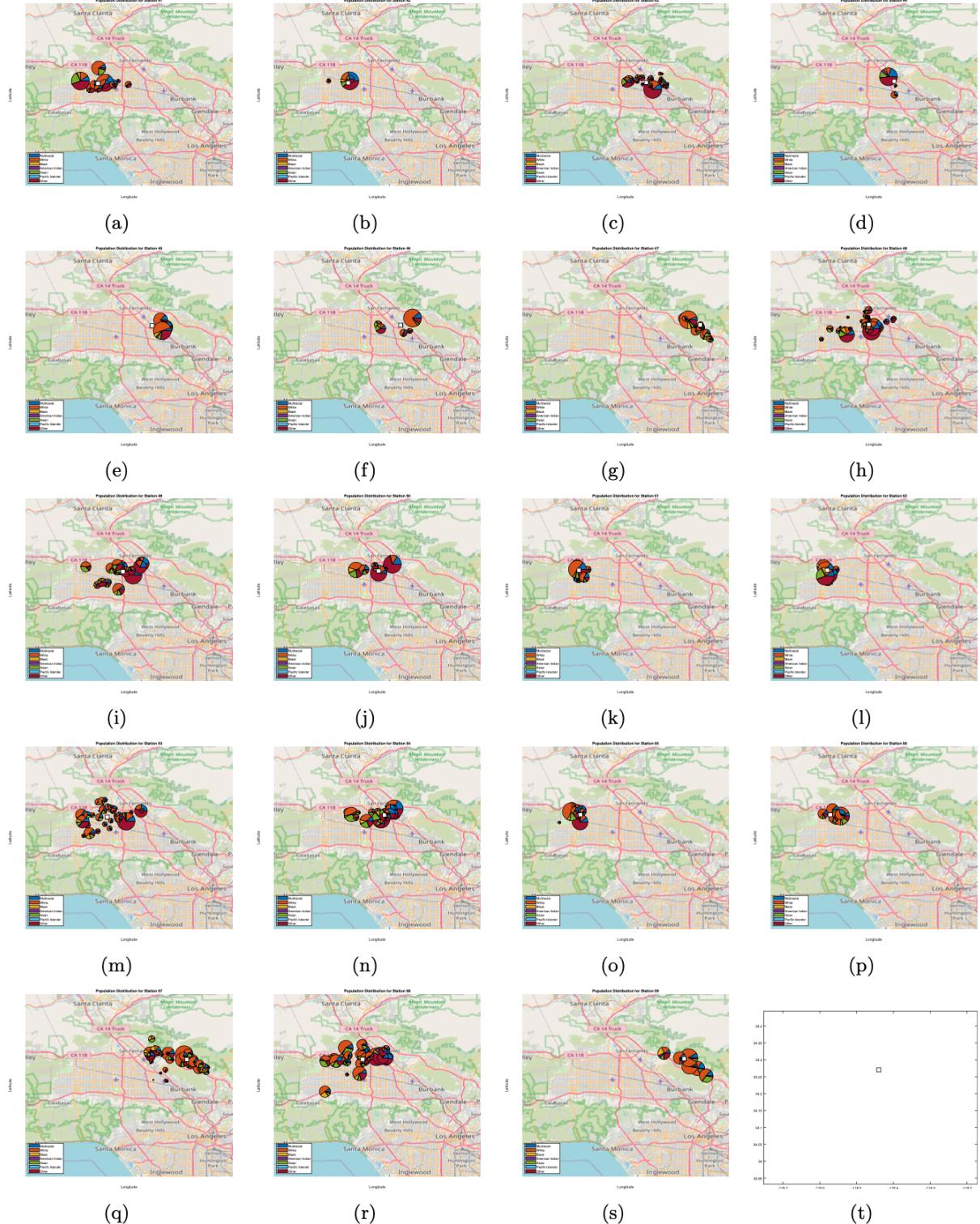
Figure 21: District boundaries for district offices 41-60 under squared distance (three groups).

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

(k)
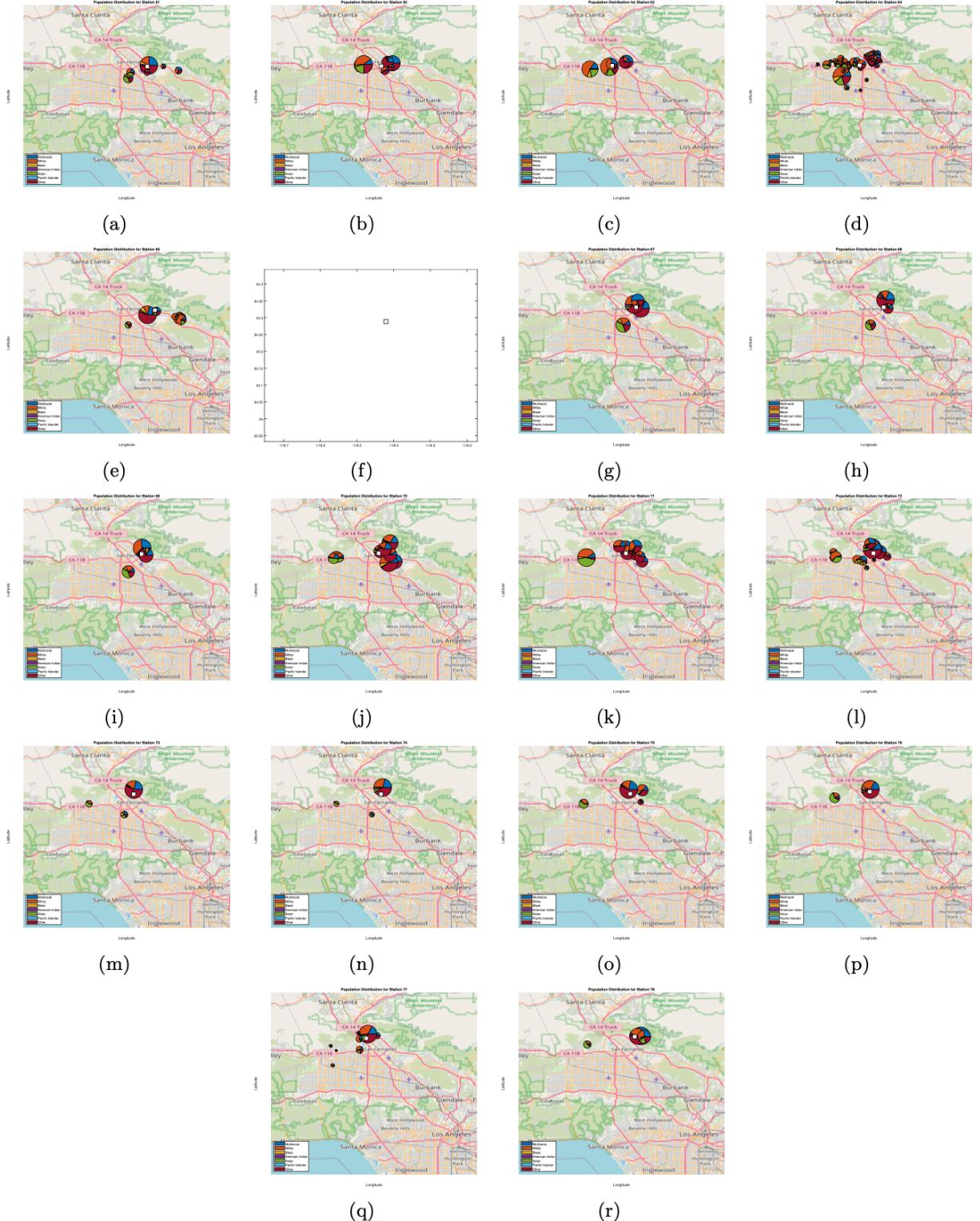

(l)

(m)

(n)

(o)

(p)

(q)

(r)

Figure 22: District boundaries for district offices 61-78 under squared distance (three groups).