

FAIRASSIGN: Stochastically Fair Driver Assignment in Gig Delivery Platforms

Daman Deep Singh
Indian Institute of Technology Delhi
New Delhi, India

Syamantak Das
IIIT Delhi
New Delhi, India

Abhijnan Chakraborty
Indian Institute of Technology Delhi
New Delhi, India

ABSTRACT

Increasing adoption of online commerce has created income opportunities for millions of delivery drivers who deliver items, from clothes and smartphones to foods and medicines, to customers. Despite their indispensability for the ecosystem, the drivers are often deprived of employment benefits and their earnings are tied to the number of successful deliveries, forcing them to go on repeated strikes to demand fair wage. In addition to low wages, there is considerable variability in driver incomes. One major component contributing to this variability is the *static assignment of drivers to delivery zones* as different zones likely have different workloads and hence earning opportunities. To reduce this variability, we directly engage with the gig delivery drivers to understand their perspectives on fair income distribution, and incorporate the same by proposing FAIRASSIGN for *dynamic assignment of drivers to delivery zones* to ensure fair distribution of earning opportunities.

Specifically, we introduce a framework for *stochastic pairwise fairness* where, based on a similarity measure between the drivers, individual drivers are assigned to probability distributions over different zones such that similar individuals are mapped to statistically similar distributions. To realize these distributions, we develop a *randomized dependent rounding* based efficient sampling algorithm such that the workload constraints in each zone are satisfied, and the expected travel cost is minimized. Extensive experiments on real-world food delivery data and semi-synthetic ecommerce data show the efficacy of FAIRASSIGN over other baselines.

CCS CONCEPTS

• **Applied computing** → *Electronic commerce; Transportation.*

KEYWORDS

Fair Driver Assignment, Last Mile Delivery, Stochastic Fairness, Dependent Rounding, Ecommerce Logistics, Food Delivery.

ACM Reference Format:

Daman Deep Singh, Syamantak Das, and Abhijnan Chakraborty. 2023. FAIRASSIGN: Stochastically Fair Driver Assignment in Gig Delivery Platforms. In *2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*, June 12–15, 2023, Chicago, IL, USA. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3593013.3594040>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

FAccT'23, June 12–15, 2023, Chicago, IL, USA

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0192-4/23/06...\$15.00

<https://doi.org/10.1145/3593013.3594040>

1 INTRODUCTION

With millions of customers resorting to online purchases for their day-to-day needs, several businesses have emerged to provide a seamless ordering and delivery experience to the customers, including food delivery platforms like Lieferando, DoorDash and Zomato [1, 52], quick commerce platforms like Zepto¹ and Gorillas, and ePharmacy (medicine delivery) platforms like Practo and 1mg [63]. Such explosive growth in online commerce has enabled skyrocketing valuations for these companies within a few years of operation [35, 54] and, at the same time, has provided earning opportunities to millions of delivery drivers² for delivering goods (from clothes and medicines to food and groceries) to the customer doorsteps, an operation formally known as *last mile delivery* [15, 56]. Although the drivers are an indispensable part of these businesses, in most scenarios (especially in the Global South), they are 'gig' workers (*freelance 'delivery partners'*) deprived of employment benefits, typically earning a small fixed fee per delivery except occasional incentives [53, 62]. Recent labor reports highlight their plight, ranging from poor working conditions to inadequate earning even after working 10+ hours a day³ [31, 42], forcing them to go on repeated strikes demanding better pay [4, 60].

In addition to lower pay, recent research works on food delivery have shown high variability in income earned by different drivers in a platform [22, 28]. Gupta et al. [28]'s detailed analysis of real food delivery data clearly shows that the income inequality is primarily caused by the difference in drivers' working areas rather than by the variability in their working hours. To counter such inequality, they propose FAIRFOODY which, instead of matching the nearest driver to a restaurant servicing an order, matches a driver with low income who can still reach the restaurant within the food preparation time [28]. In a followup work, Nair et al. [44] provide dynamic income guarantees to the drivers based on order volume and demand-supply ratio, but their proposal WORK4FOOD includes dropping drivers (i.e., not accepting app logins) during low-demand periods. While these works have a noble goal of providing *fair distribution of income*, in absence of any consultation with the actual stakeholders who are supposedly being helped – the drivers – there are risks of techno solutionism. For example, FAIRFOODY may constantly push around the drivers in different parts of the city just to get to a restaurant without carrying any food, subsequently increasing the platform cost and greenhouse emissions. Similarly in case of WORK4FOOD, having no control over when they are allowed to work further exacerbates the experiences of powerlessness of the drivers.

¹Quick commerce platforms typically promise grocery delivery in 10 minutes [21].

²In this work, we refer to the delivery agents/workers as drivers, regardless of the actual vehicle they use for delivery.

³Despite the 'gig' label, most drivers in Global South actually work full time in last-mile delivery, depending on it as their sole source of income [22].

In this work, we bridge this gap by reaching out to the delivery drivers (conduct in-person interviews) to understand the current delivery scenario and what they would consider to be a fair income distribution. Today, most quick commerce platforms (including e-groceries and ePharmacies) maintain multiple *dark stores* or *fulfillment centers (FFC)* (also known as *delivery hubs*) throughout a city. Similarly, food delivery platforms divide a city into multiple zones for operational reasons. While onboarding drivers⁴, a platform assigns a driver to a particular zone (or FFC) based on the driver's choice, geographical proximity, and approximate *capacity/vacancy* in the zone⁵ among other factors [§3]. Such static allocation of drivers to zones are explicitly announced during the onboarding process and typically do not change [16, 19]. Such *static assignments* are the root causes for the high income inequalities observed by [28, 44] since different delivery zones are likely to have different order volumes that impact the earning potential of the drivers.

We hypothesize that a *dynamic assignment* of drivers to delivery zones can lead to a fairer income distribution. However, in order to ensure fairness, we require the dynamism to be guided by a suitable notion of fairness. While the prior studies [28, 44] have focused on minimizing overall inequality, our discussion with the drivers reveal that their perceptions of (un)fairness are more rooted in their local neighborhoods. Since they do not observe the entire income distribution, especially what is happening in faraway zones in a city, their simple window to the global income distribution is through the activities of their comrades in the neighboring zones – sociological studies in other contexts have also reported similar observations [30, 34]. Based on the cues from the drivers, in this work, we attempt to reduce the *spatial inequality* [43] of incomes of drivers working in neighboring zones, albeit in an *ex-ante* manner.

Formally, we introduce a novel framework FAIRASSIGN for assigning individual drivers, embedded in a similarity space (primarily geographical but can be generalized to other notions of similarity), to *probability distributions* over the delivery zones/FFCs such that the drivers closer to each other in the similarity space are mapped to statistically similar distributions. FAIRASSIGN has two components. Firstly, we propose a measure of stochastic fairness based on the similarity space – a notion inspired by Dwork et al. [18] in the context of supervised learning problems – which allows us to formulate the dynamic assignment problem as a linear program aiming to *minimize the average travelling cost* of drivers to delivery zones/FFCs while satisfying the fairness constraints as well as additional business constraints such as bounds on the number of drivers assigned to a specific fulfillment center or delivery zone.

An optimal solution to the above linear program gives us, for each driver, a probability distribution over the delivery zones. A significant challenge is to generate the actual assignments from these distributions without violating the above constraints. In fact, any naive sampling algorithm is bound to violate such constraints. Hence, in the second step, we utilize a dependent rounding framework [23] to realize these distributions on each day such that the

expected travel cost is optimal and capacity constraints of delivery zones are satisfied. Extensive experimentation over real food delivery and semi-synthetic e-commerce datasets demonstrates the efficacy of FAIRASSIGN over several baselines.

2 BACKGROUND AND RELATED WORK

We briefly review the works on last mile delivery, food delivery, and attempts to incorporate fairness therein.

Last-mile Delivery. Due to growing urbanization and rapid expansion of online platforms, research on last-mile delivery has increased significantly. Olsson et al. [47] provides a comprehensive survey of last-mile logistics research. Since last-mile delivery spans a broad range of domains such as e-commerce, food and grocery delivery, the development of efficient delivery algorithms may necessitate consideration of domain-specific needs. For example, Escudero-Santana et al. [20] and Özarık et al. [71] looked into last-mile delivery in e-commerce; while Pan et al. [49], Weber-Snyman and Badendorst-Weiss [68] investigated the same for grocery deliveries, however without any fairness considerations.

Food Delivery. While multiple research works have attempted to find the shortest routes to customer and restaurant locations, most of these approaches make unrealistic assumptions, including apriori order arrival information [69], no consideration of road networks [57] and food preparation times [70]. FOODMATCH [32] is the state-of-the-art, being both effective and adaptable to real-world situations, but it offers no fairness guarantees. FAIRFOODY [28] is the first proposal to reduce the income inequality by modifying FOODMATCH with additional fairness objectives but introduces high additional costs for a platform. WORK4FOOD [44] provides income guarantees to drivers keeping demand-supply ratio in mind; however, it also suggests dropping drivers during low-demand periods which can significantly impact the livelihoods of delivery drivers relying solely on these platforms. Our approach differs significantly due to our focus on zone-level assignments and using a cost-efficient algorithm for individual order assignments, meeting the needs of both platforms [41] as well as the drivers [§ 3].

Stochastic Fairness. The notion of pairwise stochastic fairness was first introduced in the seminal work of Dwork et al. [18] in the context of classification problems. Recently, this idea has been extended to unsupervised learning, particularly to clustering problems [5, 12, 39]. While our proposed framework bears resemblance to [5], we deviate from it significantly to handle capacities and a lower bound on the number of assignments, which are business constraints for delivery platforms.

Procedural Fairness. While most of the works related to algorithmic fairness focus on the *fairness of outcomes*, there is a complementary line of work on *procedural fairness* (or *fairness of processes*) [26, 64]. Procedural fairness stems from the *deontological* [3] approach in moral philosophy as opposed to *consequentialism* [61], emphasizing on the justness of the decision making process and thereby ensuring fairness even when the tangible benefits are not evident. Our present work takes inspiration from the literature on procedural fairness and attempts to ensure that the driver assignments to zones follow a fair process, not directly controlling the resultant income distribution at the end.

⁴In absence of employment legalities, onboarding typically means getting background verification done, creating account in the platform app for drivers, handing over company merchandise like T-shirt and bags (often in exchange of a fee [11]) etc.

⁵Since there is no additional cost in onboarding new drivers (due to the *gig* nature of the job), the platforms often over-provision based on projected demand in a zone, which may or may not sync with the actual order volume. Ensuring ample supply of drivers help keep the delivery time low and consequently the customers happy.

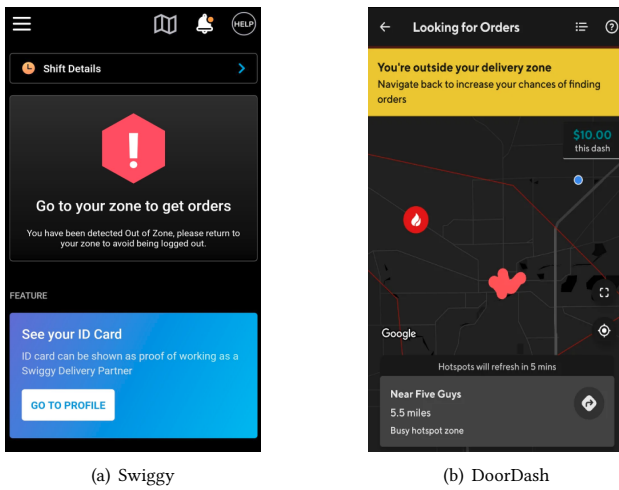


Figure 1: Screenshots from Swiggy (left) [36] and DoorDash (right) [65] delivery partner apps.

Fair Assignment. Fair division of resources has been a topic of interest in many fields, including economics [13, 51], operations research [10, 17], and computer science [7, 27, 48, 50] where the focus is on assigning items to agents (individuals or groups) fairly and efficiently where the agents have preferences (cardinal or ordinal) over the items. However, hitherto the focus has largely been on static assignments [6, 8, 37], which does not align with our aim of dynamic assignments. In contrast, studies on dynamic fair assignment have only considered one-to-one matching or disregarded practical constraints such as capacity constraints and cost effectiveness [33, 66]; hence our work complements the efforts therein.

3 FAIRNESS OF GIG DELIVERY: DRIVERS' PERSPECTIVE

In a marked departure from prior works on fair food delivery, we engage with the concerned stakeholders by conducting semi-structured interviews with the delivery drivers.

Interview Setup. We conducted in-person interviews with 30 drivers in two metro cities in India – all of them are men in their 20s and 30s – and most of them (28) are working with different food-delivery platforms. The remaining two drivers were working with a quick commerce platform. The purpose of the interviews was to know the current state of gig delivery, existence of zoning and their opinions about what would constitute a fair distribution of income opportunities. We approached the drivers when they were returning after completing deliveries in different localities, so as to not disturb their regular operations. We informed them about the purpose and requested them for a 10 minutes interview. We only proceeded with the interviews after getting informed verbal consent and we compensated their time with INR 100 which is roughly equal to their average earning per delivery (and much higher than the legal minimum wage, considering the interview duration [45]).

Key Findings. The interviews were semi-structured where we focused on eliciting narratives from the participants regarding their experiences. All of the participants reported being asked to select an available “zone” of their choice while onboarding a platform, and thereafter they could only log into the platform and pick up orders in the assigned zone. Most of the participants chose the zone closest to their residence, while a few chose a zone with more familiar restaurants. They complained about not getting orders on the way back when they deliver to customers outside their zone. Their claims can be further corroborated by screenshots posted by drivers working elsewhere (see Fig. 1).

On income disparity, they were unsure about the order volumes in farther zones; however, the majority of the drivers felt that other nearby zones with more restaurants should have higher frequency of orders. They mentioned that it is technically possible to change zones, provided there is a valid reason and “vacancy” in the desired zone, though the zone change requests are often turned down by the platforms. Concerning dynamic zone assignment, 26 out of 30 drivers had no objections as long as the zones were not too far away from their residence and the platform pays for traveling to the assigned zone, as they are already accustomed to delivering to customers in different zones.⁶ Four drivers were hesitant, as they feared that they might lose some orders while getting accustomed to new neighborhoods. Most of the drivers desired the fluidity towards seamless operation in multiple nearby zones which they feel would give them better income opportunities. However, prior information about the assigned zone before the start of a workday was deemed crucial by the drivers. They reported receiving a rating for most of the deliveries and observed that higher ratings increase the likelihood of getting more orders.

Ethical Considerations. We approached the participants with great care, not to disturb them during their jobs or when they were on move. We also reflected upon whether the questionnaire causes any mental harm. To our delight, multiple participants expressed happiness to help us as they felt empowered by being heard. We also made sure not to prime them through our questions and let the conversations free flowing. We did not record any personally identifiable information about the participants and securely stored the transcripts. We did not reveal the platform and the city names to preserve anonymity.

4 FORMALIZING THE PROBLEM OF DRIVER ASSIGNMENT

4.1 Prevalent Approach

As evident from the driver interviews, when a food delivery or quick commerce platform onboards a delivery driver, he is allotted a particular zone/FFC which becomes the area of operation and the driver needs to pickup the orders from the assigned zone/FFC every day [19]. During onboarding, the drivers get access to the delivery partner app provided by the platform which guides the drivers toward customer locations, where the routes are computed using efficient vehicle routing algorithms [9, 32, 38, 55] that consider road network alongside real-time traffic information.

⁶As mentioned earlier, they can deliver but can not pickup orders outside their assigned zone.

Formally, the task of assigning drivers to zones can be modeled as a combinatorial optimization. Suppose a set of points V represents the physical locations⁷ of drivers and another set of points C indicate the location of FFCs or zone centers. Let $d(x, y)$ denote the cost to travel between x and y where x, y are points in $V \cup C$. The function $d(x, y)$ can be used to model physical distances, time of travel, transportation cost etc. To capture the practical constraints, we also place an upper bound on the maximum number of drivers assigned to a particular zone/FFC c . We call this the *capacity* function and define as $u : C \rightarrow \mathbb{N}$. We now define the driver assignment problem.

Definition 4.1 (CAPACITATED DRIVER ASSIGNMENT). The CAPACITATED DRIVER ASSIGNMENT problem asks for an assignment $\phi : V \rightarrow C$ of every driver in V to a zone/FFC in C , such that, for every $c \in C$, $|\phi^{-1}(c)| \leq u(c)$.

Static Assignment. It is natural to design an algorithm that does a static assignment of each driver to a zone which optimizes the travel cost of the drivers, and this is the current practice followed by the platforms [19]. This problem can be formally stated as finding an assignment ϕ which minimizes the ℓ_2 -norm distance: $(\sum_{j \in V} d(j, \phi(j))^2)^{1/2}$. This can be cast as a *minimum cost b-matching problem* in bipartite graphs. The minimum cost b -matching problem reduces to the classical minimum cost flow problem [2] which has several standard efficient algorithms (see [25] for example). We call this algorithm *Minimum Cost Capacitated Assignment (MCCA)*. Once the initial zone assignment is done, actual order to driver allocation can happen through sophisticated domain specific algorithms such as FOODMATCH [32] for food delivery, or extensions of TRAVELING SALESMAN PROBLEM [9] for ecommerce deliveries. While MCCA minimizes the sum of travel cost of drivers reaching the delivery zones, it fails to accommodate any fairness considerations regarding the income distribution. If the number of orders available per driver varies a lot between zones (as observed in real food delivery data [28]), it can lead to high inequality in the earnings of different drivers.

4.2 Bringing in Fairness in Driver Assignment

Dynamic Assignment. A natural question arising out of the above discussion is - *can we do a dynamic assignment of drivers to zones/FFCs everyday to ensure lower gap in driver income?* A naive solution would be to assign a driver to different zones on different days, sequentially covering all zones in a city - this can yield a perfect balance over a long period of time. However, such an approach can blow up the travel costs for the platforms. Moreover, it can also exacerbate the experiences of powerlessness of the drivers, being pushed around to different (faraway) zones every day, completely destroying their spatial stability. Thus, the dynamism needs to be guided by a suitable notion of fairness, accepted by the actual stakeholders - the drivers.

While minimizing the overall income inequality is a noble goal, our discussion with the drivers reveal that their perceptions of (un)fairness are more rooted in their local neighborhoods. Since they do not observe the entire income distribution, especially what is happening in faraway zones in a city, their window to the global

income distribution is through the activities of their comrades in the neighboring zones. These observations are further corroborated by prior sociological studies [30, 34]. Hauser and Norton [30] showed that people rely on cues from their local environment to guess the overall income distribution and their place in it - such perceived inequalities also drive people's preferences for redistribution [30]. In this work, we attempt to reduce the *spatial inequality* [43] of incomes of drivers working in neighboring zones. In other words, we introduce a dynamic assignment framework that aims to ensure that drivers erstwhile assigned to neighboring zones (who are geographically close to each other) receive similar earning opportunities. Note that the similarity between two drivers can go beyond their geographical proximity, and include ratings or some other categorization the platform may apply. For example, many delivery platforms indeed divide drivers into multiple tiers like Blue, Bronze, Silver and Diamond, depending on their quality of service and years of association with the platforms [67]. Our underlying assumption is that if similar drivers get the chance of working on similar zones/FFCs, their incomes would be closer in the long run.

Our Notion : Stochastically Fair Assignment. We propose to split the assignment task into two phases - in the first phase, rather than doing an actual assignment, we output, for each driver, a *probability distribution* over the set of zones/FFCs. These distributions, once determined, remain static and *similar drivers* would have distributions that are *closer* to each other. However, the actual realization of the assignments is done dynamically on each day by sampling from the above probability distributions.

Let $\mathcal{F} : V \times V \rightarrow \mathbb{R}_{\geq 0}$ be a non-negative similarity measure defined over all pair of points in V . Note that \mathcal{F} is not necessarily a metric. Our algorithm works for any well-defined \mathcal{F} (such as physical distance, rating similarity, etc.). We next introduce a notion of statistical similarity between two distributions.

Definition 4.2 (Total Variational Distance). Let P, Q be two probability measures on a discrete space X . Then the total variational distance between P and Q is defined as

$$D_{TV}(P||Q) = \frac{1}{2} \sum_{x \in X} |P(x) - Q(x)|$$

Recall that the input to the problem are sets V and C denoting the driver and zone⁸ (or FFC) locations with $d(x, y)$ denoting the travel cost between points x, y . Further, $u(c), c \in C$ denotes the capacity of zone c . In order to model practical scenarios, we also consider, for each zone $c \in C$, a given lower bound $\ell(c)$ on the number of drivers that must be assigned to the zone. This ensures that no zone gets too few drivers assigned to it. However, a feasible solution might assign similar drivers to zones far away from their home locations depending on \mathcal{F} which, since the platform needs to pay for the travel to the assigned zone, might blow up platform costs and hurt the spatial stability of the drivers. Therefore, we additionally enforce that drivers are not assigned to zones too distant from their home locations.

Definition 4.3 (STOCHASTICALLY FAIR ASSIGNMENT). The STOCHASTICALLY FAIR ASSIGNMENT asks for distributions μ_v over C for each point $v \in V$ and an efficient sampling procedure from the distributions such that

⁷Home location: residential address or the place a driver usually logs into the platform to indicate the start of a work shift.

⁸We consider the 'zone center' location as a proxy for the entire zone.

Algorithm 1 FAIRASSIGN (\mathcal{I}) – Algorithm for STOCHASTICALLY FAIR ASSIGNMENT

- 1: Solve the FAIR-LP LP on instance $\mathcal{I} = (V, C, d, \mathcal{F})$ – let x^* be the optimal solution.
 - 2: Run DEPENDRND on the bipartite graph (V, C, E) with values x_{vc}^* on edge $(vc) \in E$ – let X_{vc} be the output $\{0, 1\}$ random assignments
 - 3: **for** Each $v \in V, c \in C$ **do**
 - 4: Assign driver v to c if and only if $X_{vc} = 1$
 - 5: **end for**
-

(1) The distributions satisfy the following fairness criteria

$$D_{TV}(\mu_{v_1} || \mu_{v_2}) \leq \mathcal{F}(v_1, v_2), \forall v_1, v_2 \in V \quad (1)$$

(2) For an assignment $\phi : V \rightarrow C$ obtained by above sampling from the distributions, $\forall c \in C, \ell(c) \leq |\phi^{-1}(c)| \leq u(c)$

The objective is to minimize the expected ℓ_2 -cost, that is,

$$\left(\sum_{v \in V} \mathbb{E}_{c \sim \mu_v} (d(v, c)^2) \right)^{\frac{1}{2}}.$$

Remark: We would like to emphasize that all bounds guaranteed by our algorithm in §5 extends to all f-divergence measures and any ℓ_p -norm of cost for any $p \geq 1$ (In equation (1)). However, the specific choice of total variation distance D_{TV} and ℓ_2 -norm of cost is motivated by practical considerations of computability. Specifically, as shown in §5, the above formulation allows us to formulate a *linear program* as one of the crucial steps of our algorithm for which very efficient solvers exist.

5 FAIRASSIGN: AN ALGORITHM FOR STOCHASTICALLY FAIR DRIVER ASSIGNMENT

We present a novel two-phased algorithm FAIRASSIGN (Algorithm 1) to solve STOCHASTICALLY FAIR ASSIGNMENT defined earlier.

Phase 1. The first phase consists of generating the distributions μ_v for each $v \in V$ as required in STOCHASTICALLY FAIR ASSIGNMENT. To efficiently solve this, we develop a linear program to output the distributions with all the desired properties of Definition 4.3. The LP on the instance $\mathcal{I} = (V, C, d, \mathcal{F})$ is given as

$$\text{FAIR-LP } (\mathcal{I}) : \min \sum_{v \in V} \sum_{c \in C} x_{vc} d(v, c)^2 \quad (2)$$

$$\text{s.t. } \sum_{c \in C} x_{vc} = 1, \forall v \in V \quad (3)$$

$$D_{TV}(\vec{x}_{v_1} || \vec{x}_{v_2}) \leq \mathcal{F}(v_1, v_2), \forall v_1, v_2 \in V \quad (4)$$

$$0 \leq x_{vc} \leq 1, \forall v \in V, c \in C \quad (5)$$

$$\sum_{v \in V} x_{vc} \geq \ell(c), \forall c \in C \quad (6)$$

$$\sum_{v \in V} x_{vc} \leq u(c), \forall c \in C \quad (7)$$

Once we have an optimal solution to FAIR-LP (\mathcal{I}), $\forall v \in V$ and $c \in C$, we can think of x_{vc} as the probability that the client v is assigned to the zone c . Hence, \vec{x}_v will give the desired distribution μ_v corresponding to driver v over the set of zones C . Constraint (3) simply enforces that x_{vc} for a fixed v indeed forms a probability

Algorithm 2 DEPENDRND(V, C, E, x) – Dependent Rounding Algorithm

- 1: Let $F = \{(vc) \in E : 0 < x_{vc} < 1\}$
 - 2: **while** $|F| > 0$ **do**
 - 3: Find a simple cycle or a *maximal* path P in the subgraph (V, C, F) using DFS and decompose it in to two disjoint *matchings* M_1 and M_2
 - 4: $\alpha \leftarrow \min\{\varepsilon > 0 : \exists(vc \in M_1) : x_{vc} + \varepsilon = 1 \vee \varepsilon > 0 : \exists(vc \in M_2) : x_{vc} - \varepsilon = 0\}$
 - 5: $\beta \leftarrow \min\{\varepsilon > 0 : \exists(vc \in M_1) : x_{vc} - \varepsilon = 0 \vee \varepsilon > 0 : \exists(vc \in M_2) : x_{vc} + \varepsilon = 1\}$
 - 6: With probability $\beta/(\alpha + \beta)$, set $x_{vc} \leftarrow x_{vc} + \alpha, \forall(vc) \in M_1, x_{vc} \leftarrow x_{vc} - \alpha, \forall(vc) \in M_2$
 - 7: With the complementary probability $\alpha/(\alpha + \beta)$, set $x_{vc} \leftarrow x_{vc} - \beta, \forall(vc) \in M_1, x_{vc} \leftarrow x_{vc} + \beta, \forall(vc) \in M_2$
 - 8: **end while**
 - 9: **for all** $(vc) \in E$ **do**
 - 10: $X_{vc} \leftarrow x_{vc}$
 - 11: **end for**
 - 12: Return the assignments X_{vc} , for all $(vc) \in E$
-

distribution. The constraint (4) ensures stochastic fairness. Constraints (6) and (7) enforce the lower bounds and upper bounds (or capacity) constraints on the zones respectively. An optimal solution to FAIR-LP can be found in time polynomial in $|V|, |C|$.

Phase 2. Now we describe the sampling from the distributions $\mu_v, \forall v \in V$ obtained by solving FAIR-LP. Let x^* be an optimal solution to FAIR-LP. A natural approach is to carry out an independent sampling for each driver $v \in V$ - that is assign driver v to FFC c with a probability x_{vc}^* . However, in such a sampling we cannot guarantee that the number of drivers assigned to a specific zone will lie between $\ell(c), u(c)$ with probability 1. Take for example a simple instance with 4 drivers and 2 zones with the capacity of each zone being 2. Suppose each driver is assigned to each of the two zones with a probability 1/2 by FAIR-LP. Now in the naive sampling procedure, with probability 1/16 (since the sampling is done independently for each driver), all 4 drivers will be assigned to the same zone leading to an infeasible assignment.

In order to overcome this challenge, we resort to *dependent rounding* [23]. To understand the framework, think of the problem setup of STOCHASTICALLY FAIR ASSIGNMENT as a bipartite graph (V, C, E) - the driver and zone locations form the bipartition here. Further, $d(v, c)$ denotes the cost of each edge $(vc) \in E, v \in V, c \in C$. Now upon solving FAIR-LP, we obtain a fractional value $x_{vc}^* \in [0, 1]$ for every edge (vc) . An actual assignment would now correspond to rounding each value x_{vc}^* to a random variable $X_{vc} \in \{0, 1\}$ in such a way that the following properties hold.

- **(P1) : Marginal distribution.** For every $v \in V, c \in C$, $\Pr[X_{vc} = 1] = x_{vc}^*$
- **(P2) : Degree preservation.** For any $i \in V \cup C$, define $d_i = \sum_{j:(ij) \in E} x_{ij}^*$ as the fractional degree of i in x^* and $D_i = \sum_j X_{ij}$ as its internal degree in the rounded solution. Then we must have $D_i \in \{d_i, \lceil d_i \rceil\}$. In particular, if d_i is an integer, this ensures $D_i = d_i$ with probability 1.

(P1) is essential to ensure that the expected cost of the assignment is equal to the cost of the optimal solution x^* to FAIR-LP. Property (P2) is crucial to preserve the capacity and lower bound constraints as required in the definition of STOCHASTICALLY FAIR ASSIGNMENT. Gandhi et al. [23] (Section 2.1) describes a randomized algorithm running in time $O((|V|+|C|) \cdot |E|)$ for the dependent rounding problem satisfying properties (P1), (P2). We call this algorithm DEPENDRND and describe it as Algorithm 2 for the sake of completeness. We apply this sampling directly on each day to output the final assignments of drivers to zones. This concludes the description of our algorithm. We summarize the main result through the following theorem.

THEOREM 5.1. *The algorithm FAIRASSIGN outputs probability distributions for each driver $v \in V$ over the set of zones C that satisfy the stochastic fairness constraints (4.3). Further, the subroutine DEPENDRND efficiently outputs a randomized assignment of drivers to zones such that the expected ℓ_2 -norm cost of this assignment is optimal and all the capacity and lower bound constraints for each zone are satisfied.*

PROOF. The first property just follows from the fact that FAIR-LP is a correct linear program for STOCHASTICALLY FAIR ASSIGNMENT as shown previously.

Now, the expected ℓ_2 -norm cost of assignment is

$$\begin{aligned} & \sum_{v \in V} \mathbb{E}_{c \sim \mu_v} (d(v, c)^2) \\ &= \sum_{v \in V} \sum_{c \in C} \Pr[X_{vc} = 1] (d(v, c)^2) \\ &= \sum_{v \in V} \sum_{c \in C} x_{vc}^* (d(v, c)^2) \end{aligned}$$

The last equality follows from (P1) of DEPENDRND and further, the last term is exactly equal to the optimal solution to FAIR-LP. Finally (P2) ensures the desired capacity and lower bounds for each zone in C . \square

Applying FAIRASSIGN in the wild. FAIRASSIGN performs zone-level assignment. For further order-level assignments and routing, we need to use domain-specific last-mile algorithms. For instance, in the food-delivery domain, an algorithm such as FOODMATCH [32] can handle both order-level assignment and routing. For e-commerce delivery and similar settings, we can employ some variant of TRAVELING SALESMAN PROBLEM [9]. Note that TRAVELING SALESMAN PROBLEM can restrict a driver to deliver orders within a statically assigned zone. While FOODMATCH theoretically has no such zone-level restrictions, empirically we found most of the orders for a driver to be contained within their base zones.

6 EXPERIMENTAL EVALUATION

We conduct extensive experimentation using a machine with Intel Xeon CPU @ 2.10GHz and 252GB RAM, running Ubuntu 18.04.3 LTS. All algorithms are implemented in Python 3.9. We use IBM CPLEX [14] and Gurobi Optimizer [29] for solving the linear program FAIR-LP. We present the experimentation details including the food delivery dataset and the baselines, followed by the detailed results. We further demonstrate the generalizability of our framework by applying it to a different domain – e-commerce delivery.

6.1 Food-delivery Dataset

We use nine days of real-world food delivery data from 3 major Indian cities [28, 32]; we refer to the cities as City A , B and C . The dataset consists of the trajectories of the delivery vehicles, the road network of the city, and metadata such as vehicle IDs, order information, locations of restaurants and customers, zone boundary locations, etc. The platform segments the city into delivery zones, each corresponding to a prominent neighborhood.

We determine the capacity of each zone using data from the first three days, where the minimum and maximum capacity of the zones are set as 30% and 100% of the average number of drivers active in the respective zones. We use the remaining 6 days of data for experiments which includes both weekdays and weekends. We observe that the drivers log into the platform from similar locations every day, hence we consider the *home* (or *base*) location of a driver as the location from where he has logged in the first day. We compute the “zone center” as the arithmetic mean of the zone boundary locations provided in the dataset. Table 1 summarizes the dataset characteristics.

6.2 Baselines

We compare the performance of FAIRASSIGN against the following baseline methods:

- **FOODMATCH:** An efficient last-mile delivery algorithm for the food-delivery domain [32].
- **FAIRFOODY:** A fair food delivery algorithm [28], aimed at reducing the income gap between drivers.
- **Least Income Priority Assignment (LIPA):** Everyday this method greedily assigns the drivers with hitherto lowest incomes to zones with higher order volumes, subject to the capacity constraints.
- **Round Robin Assignment (ROUNDROBIN):** This strategy assigns the delivery drivers to zones in a round-robin manner. This process will be repeated as long as the capacity constraints are satisfied.

6.3 Driver Similarity Measures

We consider the following two similarity measures between drivers:

- **Physical Distance (\mathcal{F}_1):** The euclidean distance between the latitude-longitude level locations of the drivers. This measure follows the locally-rooted perception of fairness of the drivers, as seen in §3. For easier operationalization, we limit the fairness criteria (equation 1) to apply only to drivers within a specified distance of each other. This distance can be a hyperparameter chosen independently for every scenario. Imposing this restriction also benefits computationally by limiting the number of pairwise comparisons.
- **Weighted Distance-Rating measure (\mathcal{F}_2):** A linear combination of *physical distance* (\mathcal{F}_1) and *rating*. The rating serves as a proxy for any performance measure. For our experimentation, we define the rating for each driver as a number in the interval $[0, 5]$ with step size 0.1, sampled from a positively-skewed Gaussian distribution with mean 3.5 and standard deviation 1. The choice of the distribution parameters is based on rating distributions observed in real-world datasets [40].

$$\mathcal{F}_2 = w_1 \cdot \mathcal{F}_1 + w_2 \cdot \text{rating} \quad (8)$$

Location	#Orders	#Drivers	#Zones
A	23,442	2,454	21
B	159,160	13,429	44
C	112,745	10,608	49

Table 1: Statistics of the food-delivery dataset.

6.4 Evaluation Metrics

Fairness metrics. To evaluate the fairness of driver assignments, we use the following three metrics:

- *Gini Index*: Gini captures the *relative mean absolute difference* of income between any pair of drivers [24], and can be measured as:

$$\text{Gini} = \frac{\sum_{i=1}^N \sum_{j=1}^N |y_i - y_j|}{2N \sum_{i=1}^N y_i} \quad (9)$$

where y_i is the income of driver i , and N is the number of drivers. Gini ranges between 0 to 1, with 0 denoting perfect equality.

- *Spatial Inequality Index*: Mota et al. [43] proposed a variation of Gini, where instead of comparing the income of any pair of drivers, we only consider the drivers who are located within a certain distance from each other, and then aggregate over all drivers.

$$\text{Spat. Ind.} = \frac{\sum_{i=1}^N \frac{1}{N_i} \sum_{j=1}^{N_i} |y_i - y_j|}{\sum_{i=1}^N y_i} \quad (10)$$

where N_i is the number of drivers comparable to i .

- *Income Gap*: This metric captures the income disparity between any two drivers per unit difference in the similarity measure under consideration. It represents the idea that it would be fine if two dissimilar drivers earn differently, but similar drivers should have similar incomes.

$$\text{Inc. Gap} = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{|y_i - y_j|}{\text{distance}(i, j)} \quad (11)$$

Spatial Stability. In dynamic assignment, a driver is assigned a zone every day using FAIRASSIGN. We want to measure how frequently the assigned zone changes. Let L_d be the list of zones assigned to a driver d over the span of N days. We define the spatial stability of driver d as:

$$\text{Spatial Stability}(L_d) = H \times R \quad (12)$$

where H is the entropy of the frequency distribution of distinct zones in L_d , and R is the number of zone changes in L_d i.e., the number of times a driver receives a new zone assignment during consecutive days. For example, the list $L_d = [1, 2, 2, 3, 3, 3, 2, 2]$ has $R = 3$. In essence, H captures the spread of the assignments in L_d , whereas R quantifies the extent of contiguous assignments to the same zone. The choice to use R stems from the fact that uninterrupted stretches of assignment to the same zone promote more spatial stability.

We further define the spatial stability of an assignment algorithm as the mean of the spatial stability of all the drivers. Lower values

Location	#Orders	#Drivers	#Zones
Minas Gerais (MG)	11,345	92	10
Rio de Janeiro (RJ)	12,337	123	10
São Paulo (SP)	40,480	350	10

Table 2: Statistics of the e-commerce dataset.

indicate more spatial stability. Naturally, static assignments are the most spatially stable with a spatial stability metric value of 0. In contrast, round-robin assignments are the most spatially unstable.⁹

Cost to platform. The cost of driver assignment is determined by combining the *first-mile* and *last-mile* distances, computed over all drivers and all days. The *first-mile* distance for a particular driver refers to the distance traveled to reach the designated zone, while the *last-mile* distance represents the total distance the driver travels throughout the day to deliver the orders assigned to them. Since a delivery platform pays the drivers to compensate for their travels (refer to §3 for details), we consider the average distance across all drivers as a proxy for the “platform cost”. A platform will prefer a scheme with a lower cost.

6.5 Results

As explained in §5, we first run different assignment baselines for assigning drivers to zones and then apply FOODMATCH to simulate the last-mile delivery. Furthermore, to maintain the spatial stability of the drivers and lower the platform costs, we enforce that drivers can only be assigned to K nearest zones to their home locations. The results have been presented for $K=10$, $w_1=0.6$ and $w_2=0.4$. We obtain similar results for other viable combinations of K , w_1 , and w_2 , however, they have been excluded for brevity.

Tables 3 and 4 show the performance of different algorithms with \mathcal{F}_1 and \mathcal{F}_2 similarity measures respectively. The rows for FOODMATCH and FAIRFOODY reflect the results obtained on applying these delivery algorithms as is on the dataset, without any alteration in the assignment on top of the static assignment inherently present in the dataset¹⁰. The other results are obtained on the application of FOODMATCH after the indicated assignment baseline. For both the \mathcal{F}_1 and \mathcal{F}_2 similarity measures, we see that FAIRFOODY outperforms all algorithms in terms of fairness in most cases but incurs a significantly higher cost. FAIRASSIGN, on the other hand, incurs only a slight increase in cost compared to FOODMATCH while achieving a considerable improvement in all fairness metrics. Note that the rating space can differ significantly from the physical distance space, so the cost-fairness trade-off for \mathcal{F}_2 might not be as impressive as for \mathcal{F}_1 . Nevertheless, the \mathcal{F}_2 has its own merit as it allows the platforms the flexibility to factor in a driver’s performance while assigning zones. Conclusively, these results suggest that the application of FAIRASSIGN for zone-level assignment prior to an efficient last-mile delivery algorithm can lead to a fairer income distribution while incurring a considerably lower cost than a purely

⁹However, *capacitated* ROUNDROBIN, as described in §6.2, may not necessarily perform the worst in terms of spatial stability.

¹⁰Such static assignment arises due to real-world hiring processes as explained in the introduction.

Algorithm	A					B					C				
	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.
FAIRASSIGN	<u>0.302</u>	<u>0.755</u>	<u>0.282</u>	<u>0.486</u>	4.314	0.162	<u>1.378</u>	0.299	0.228	<u>2.055</u>	<u>0.197</u>	<u>1.031</u>	<u>0.393</u>	<u>0.350</u>	<u>0.266</u>
FOODMATCH	<i>0.370</i>	<i>0.732</i>	<i>0.359</i>	<i>0.552</i>	0.0	<i>0.210</i>	1.246	0.362	<i>0.312</i>	0.0	<i>0.218</i>	0.990	0.434	<i>0.398</i>	0.0
FAIRFOODY	0.137	<i>0.988</i>	0.133	0.234	0.0	0.160	<i>1.625</i>	<i>0.514</i>	0.232	0.0	0.124	<i>1.227</i>	0.339	0.231	0.0
LIPA	0.311	0.771	0.312	0.499	<u>2.147</u>	<u>0.147</u>	1.608	0.326	0.215	6.678	0.202	1.094	0.468	0.372	2.743
ROUNDROBIN	0.305	0.795	0.296	0.489	<i>6.969</i>	0.140	1.619	<u>0.315</u>	0.250	<i>7.127</i>	0.209	1.171	<i>0.478</i>	0.372	<i>2.922</i>

Table 3: Performance of different algorithms on food delivery dataset with \mathcal{F}_1 . The best results are highlighted in blue and emboldened, while the second-best results are underlined. The worst results are indicated in red and italicized.

Algorithm	A					B					C				
	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.
FAIRASSIGN	0.295	0.765	<u>0.560</u>	<u>0.459</u>	<u>0.727</u>	0.197	1.234	1.385	0.386	<u>0.313</u>	<u>0.198</u>	<u>1.043</u>	<u>1.025</u>	<u>0.394</u>	<u>2.271</u>
FOODMATCH	<i>0.370</i>	<i>0.732</i>	<i>0.714</i>	<i>0.544</i>	0.0	<i>0.210</i>	1.246	1.528	<i>0.404</i>	0.0	<i>0.218</i>	0.990	1.044	<i>0.433</i>	0.0
FAIRFOODY	0.138	<i>0.988</i>	0.377	0.248	0.0	0.160	<i>1.625</i>	<i>2.205</i>	0.305	0.0	0.124	<i>1.227</i>	0.766	0.240	0.0
LIPA	0.311	<u>0.747</u>	0.601	0.484	2.147	<u>0.147</u>	1.608	<u>1.375</u>	<u>0.292</u>	6.678	0.202	1.094	1.080	0.403	2.743
ROUNDROBIN	0.305	0.795	0.593	0.476	<i>6.969</i>	0.140	1.619	1.335	0.278	<i>7.127</i>	0.209	1.171	<i>1.185</i>	0.407	<i>2.922</i>

Table 4: Performance of different algorithms on food delivery dataset with \mathcal{F}_2 . The best results are highlighted in blue and emboldened, while the second-best results are underlined. The worst results are indicated in red and italicized.

order fairness-focused approach hence maintaining a nice platform cost-fairness trade-off.

6.6 Generalizing beyond food delivery

To show the generalizability of FAIRASSIGN, we conduct experimentation on a dataset from the ecommerce domain.

Ecommerce dataset. The highly competitive nature of ecommerce makes customer, delivery driver, and fulfillment center information highly confidential business assets¹¹, and hence such information is rarely available. To circumvent this difficulty, we generate a part-real-part-synthetic dataset using a public dataset from the Brazilian ecommerce website `olist.com` [46]. The dataset contains information about 100,000 random orders placed between 2016 to 2018 in multiple states across Brazil. It contains the customer locations, seller locations and order information, including timestamps. We consider the top 3 states as per the order volume, which account for nearly 65% of the dataset. Due to the sampling of orders, the number of orders in a single day was low. Hence, we grouped the orders for 15 days together and considered the time period to be a single day.

We use the initial 20% of this dataset (sorted in chronological order) to generate the location of fulfillment centers (FFCs). This is equivalent to observing customer demands during the initial period of launching a service and accordingly planning the logistics. We use k -median uncapacitated facility location algorithm [59] on the customer locations to generate the locations of FFCs. We generate the driver locations by dividing the states into $M \times N$ grids, calculating the number of customers in each grid, multiplying it with a random number from $[0.5, 1.5]$ to get the number of delivery drivers and then assigning their locations uniformly randomly in the grid. Finally, since the FFCs can only accommodate a limited number of delivery drivers, we calculate the 70-th percentile of the daily orders for each center (by assigning an order to its closest center) and set the minimum and maximum capacity as 30% and 100%

¹¹Sahay et al. [58] reported that as soon as the competitors figure out the presence of one company’s fulfillment center in an area, they set up their own centers in the vicinity, possibly anticipating high customer demand.

of this order quantity. Note that k and grid size $M \times N$ are tunable hyperparameters. Table 2 summarizes the dataset characteristics.

Baselines. In the case of ecommerce, we consider the following baseline methods:

- *Minimum Cost Capacitated Assignment (MCCA)*: described in [§ 4.1]
- *Minimum Cost Capacitated Assignment with Lower Bound (MCCA-L)*: An extension of MCCA where alongside capacity, we also consider a minimum number of drivers needed to keep an FFC running. We use the min-cost-max-flow algorithm with two constraints – upper bound and lower bound – to assign drivers to centers.
- *Least Income Priority Assignment (LIPA)*: described in [§ 6.2]
- *Round Robin Assignment (ROUNDROBIN)*: described in [§ 6.2]

Results. For each state, we first run a particular algorithm (a baseline method or FAIRASSIGN) to assign drivers to fulfillment centres (FFCs). Since the number of FFCs and grid sizes are parameters in our dataset, we have varied K from 5 to 20 and grid size from 5×5 to 15×15 in different experiments. For the experiments corresponding to \mathcal{F}_2 similarity measure, we consider $w_1=0.7$ and $w_2=0.3$. We present the results for 10 FFCs and 7×7 grid size. Other parameter combinations give qualitatively similar results but have been omitted for brevity. Since we employ the same last-mile delivery algorithm [9] after each assignment algorithm, so we report the average ‘first-mile’ distance as the platform cost. Note that the results presented here are the average of 10 experimental trials.

Table 5 shows the performance of different algorithms in all three states while considering the \mathcal{F}_1 similarity measure. We observe that the de facto algorithm for cost (or average distance) minimization, – MCCA-L indeed performs best in terms of cost but performs poorly on all fairness metrics. LIPA provides the fairest assignment possible, but at the cost of a large increase in distance. ROUNDROBIN, in some cases, performs slightly better than FAIRASSIGN in terms of fairness but it does so at an unreasonably higher cost. Our proposed algorithm FAIRASSIGN achieves a nice trade-off between cost and fairness, being closer to MCCA and MCCA-L in cost and closer to LIPA in terms of fairness metrics.

Algorithm	RJ					SP					MG				
	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.
FAIRASSIGN	<u>0.124</u>	0.221	1.184	0.110	<u>15.205</u>	0.096	0.560	0.135	0.070	<u>6.933</u>	<u>0.119</u>	0.802	0.124	0.130	<u>3.9183</u>
MCCA	<i>0.268</i>	0.160	<i>2.870</i>	<i>0.344</i>	0.0	0.122	0.511	0.187	0.106	0.0	0.134	0.762	0.139	0.138	0.0
MCCA-L	0.241	<u>0.170</u>	2.750	0.336	0.0	<i>0.123</i>	<u>0.511</u>	<i>0.188</i>	<i>0.107</i>	0.0	<i>0.138</i>	<u>0.766</u>	<i>0.147</i>	<i>0.153</i>	0.0
LIPA	0.127	0.425	1.349	0.229	33.241	0.012	1.294	0.022	0.020	<i>27.787</i>	0.021	2.299	0.023	0.039	25.231
ROUNDROBIN	0.084	<i>0.820</i>	1.430	<u>0.129</u>	<i>34.108</i>	<u>0.079</u>	<i>1.744</i>	<u>0.111</u>	<u>0.058</u>	24.078	0.120	<i>2.781</i>	<u>0.111</u>	<u>0.093</u>	<i>44.582</i>

Table 5: Performance of different algorithms on e-commerce dataset with \mathcal{F}_1 . The best results are highlighted in blue and emboldened, while the second-best results are underlined. The worst results are indicated in red and italicized.

Algorithm	RJ					SP					MG				
	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.	Gini	Avg Dist.	Inc. Gap	Spat. Ind.	Spat. St.
FAIRASSIGN	<u>0.092</u>	0.379	0.372	0.118	<u>26.942</u>	<u>0.070</u>	0.885	<u>0.060</u>	0.038	<u>36.523</u>	<u>0.110</u>	1.593	0.092	<u>0.038</u>	27.807
MCCA	<i>0.268</i>	0.160	<i>1.148</i>	<i>0.378</i>	0.0	0.122	0.511	0.148	0.042	0.0	0.134	0.762	0.131	0.046	0.0
MCCA-L	0.241	<u>0.170</u>	1.050	0.358	0.0	<i>0.123</i>	<u>0.511</u>	<i>0.149</i>	0.042	0.0	<i>0.138</i>	<u>0.766</u>	<i>0.138</i>	0.049	0.0
LIPA	0.143	0.421	0.442	0.257	<i>30.2189</i>	0.023	1.280	0.016	0.019	<i>37.039</i>	0.021	2.299	0.016	0.016	27.339
ROUNDROBIN	0.084	<i>0.820</i>	<u>0.426</u>	<u>0.127</u>	26.958	0.079	<i>1.744</i>	0.084	0.058	<i>50.737</i>	0.120	<i>2.781</i>	0.104	0.039	<i>42.276</i>

Table 6: Performance of different algorithms on e-commerce dataset with \mathcal{F}_2 . The best results are highlighted in blue and emboldened, while the second-best results are underlined. The worst results are indicated in red and italicized.

	Algorithm	Food Delivery			ecommerce		
		A	B	C	RJ	SP	MG
Total Cost	FOODMATCH/ MCCA-L (\mathcal{F}_1 & \mathcal{F}_2)	0.732	1.246	0.990	2.987	1.740	5.925
	FAIRASSIGN (\mathcal{F}_1)	0.755	1.378	1.031	2.900	1.785	5.977
	FAIRASSIGN (\mathcal{F}_2)	0.765	1.234	1.043	3.002	2.093	6.695
		<i>(+2.9%)</i>	<i>(+10.59%)</i>	<i>(+4.1%)</i>	<i>(-2.9%)</i>	<i>(+2.6%)</i>	<i>(+5.2%)</i>
		<i>(+4.4%)</i>	<i>(-0.96%)</i>	<i>(+5.3%)</i>	<i>(+0.5%)</i>	<i>(+3.5%)</i>	<i>(+13%)</i>
Income Gap	FOODMATCH/ MCCA-L (\mathcal{F}_1)	0.359	0.362	0.434	2.750	0.188	0.147
	FAIRASSIGN (\mathcal{F}_1)	0.282	0.299	0.393	1.184	0.135	0.124
		<i>(-21.44%)</i>	<i>(-17.4%)</i>	<i>(-9.45%)</i>	<i>(-56.9%)</i>	<i>(-28.2%)</i>	<i>(-15.6%)</i>
	FOODMATCH/ MCCA-L (\mathcal{F}_2)	0.714	1.523	1.044	1.050	0.149	0.138
	FAIRASSIGN (\mathcal{F}_2)	0.560	1.385	1.025	0.372	0.060	0.092
		<i>(-21.6%)</i>	<i>(-9.06%)</i>	<i>(-1.8%)</i>	<i>(-64.6%)</i>	<i>(-59.7%)</i>	<i>(-33.3%)</i>

Table 7: Cost vs Fairness due to MCCA-L, FOODMATCH and FAIRASSIGN with \mathcal{F}_1 and \mathcal{F}_2 similarity measures. The percentage differences in total cost and income gap between FAIRASSIGN vs MCCA or FOODMATCH are shown in parenthesis. The improvement due to FAIRASSIGN is highlighted in blue (with a - sign), while declines are indicated in red (with a + sign).

The results corresponding to the \mathcal{F}_2 similarity measure in table 6 show performance trends similar to those observed for \mathcal{F}_1 similarity measure. Here we see that FAIRASSIGN outperforms most baselines, never falling below second best in terms of fairness with a reasonable increase in cost. The results demonstrate the versatility of FAIRASSIGN in handling a blend of various similarity measures, providing flexibility in creating domain-specific metrics.

6.7 Cost-fairness trade-off of FAIRASSIGN

In this discussion, we show a more fine-grained comparison of FAIRASSIGN with a cost-efficient algorithm relevant to each domain – FOODMATCH for food delivery and MCCA-L for ecommerce (followed by the application of a last-mile delivery algorithm [9]). Specifically, we show, as depicted in table 7, the change in total cost averaged over all drivers and all days versus the income gap, where the total cost is the sum of the first-mile distance and the last-mile distance [§ 6.4].

We observe that, compared to MCCA-L or FOODMATCH, applying FAIRASSIGN results in only a slight increase in overall cost while significantly reducing the income gap. These results indicate that incorporating FAIRASSIGN can be a cost-effective means of promoting fairness in last-mile delivery systems.

7 CONCLUSION

In this paper, we proposed a dynamic assignment algorithm – FAIRASSIGN – to assign delivery drivers to delivery zones or fulfillment centers such that two drivers operating in nearby zones get similar earning opportunities. We engaged directly with the delivery drivers to design the algorithm in a procedurally fair manner. We recognize that the perception of (un)fairness among drivers, which is rooted in their local context, may be attributable to a dearth of global information, such as the average income earned per delivery zone, and the lack of transparency in the platforms. Nonetheless, our present work is concerned with the gig delivery ecosystem’s current and foreseeable future state. We demonstrated the general applicability of our framework via extensive experimentation on datasets from two distinct domains – food delivery and ecommerce, and found that we can gain significantly in terms of fairness with minimal increase in the additional travel cost. In this work, we considered same delivery fee per order, and hence income is directly proportional to number of deliveries. In future, we plan to extend our work to include scenarios with variable payments based on size or weight of the delivered items.

Reproducibility. The algorithms and codebase are available at <https://github.com/ddsb01/FairAssign>.

ACKNOWLEDGMENTS

We would like to express our sincere gratitude to the anonymous reviewers for their valuable feedback which helped in improving the quality of the paper. We also extend our appreciation to our colleagues Kshitiz Jain, Devesh Pant and Kunal Dargan for helping us in conducting the interviews with the delivery drivers. Additionally, we thank Piyush Gupta and Nihesh Anderson for their assistance during the early stages of this work.

REFERENCES

- [1] Kabir Ahuja, Vishwa Chandra, Victoria Lord, and Curtis Peens. 2021. Ordering in: The rapid evolution of food delivery.
- [2] Ravindra K Ahuja, Thomas L Magnanti, and James B Orlin. 1993. Network Flows: Theory, Algorithms, and Applications.
- [3] Larry Alexander and Michael Moore. 2021. Deontological Ethics. In *The Stanford Encyclopedia of Philosophy* (Winter 2021 ed.), Edward N. Zalta (Ed.). Metaphysics Research Lab, Stanford University.
- [4] Linah Alsaafin. 2022. Dubai: Deliveroo riders strike for better pay and work conditions. <https://www.aljazeera.com/news/2022/5/2/dubai-deliveroo-riders-strike-for-better-pay-and-work-conditions>.
- [5] Nihesh Anderson, Suman K Bera, Syamantak Das, and Yang Liu. 2020. Distributional individual fairness in clustering.
- [6] Haris Aziz, Serge Gaspers, Simon Mackenzie, and Toby Walsh. 2015. Fair assignment of indivisible objects under ordinal preferences. *Artificial Intelligence* 227 (2015), 71–92. <https://doi.org/10.1016/j.artint.2015.06.002>
- [7] Haris Aziz and Simon Mackenzie. 2017. A Discrete and Bounded Envy-Free Cake Cutting Protocol for Any Number of Agents. arXiv:1604.03655 [cs.DS]
- [8] Siddharth Barman and Sanath Kumar Krishna Murthy. 2017. Approximation Algorithms for Maximin Fair Division. *CoRR abs/1703.01851* (2017). arXiv:1703.01851 <http://arxiv.org/abs/1703.01851>
- [9] Richard Bellman. 1962. Dynamic Programming Treatment of the Travelling Salesman Problem. *J. ACM* 9, 1 (jan 1962), 61–63. <https://doi.org/10.1145/321105.321111>
- [10] Dimitris Bertsimas, Vivek Farias, and Nikolaos Trichakis. 2011. Fairness, Efficiency and Flexibility in Organ Allocation for Kidney Transplantation. *Operations Research* 61 (02 2011). <https://doi.org/10.1287/opre.1120.1138>
- [11] blinkit. 2023. Blinkit delivery partner. <https://blinkit.com/delivery>.
- [12] B Brubach, D Chakrabarti, J Dickerson, A Srinivasan, and L Tsepenekas. 2021. Fairness, Semi-Supervised Learning, and More: A General Framework for Clustering with Stochastic Pairwise Constraints. <https://arxiv.org/abs/2103.02013>
- [13] Eric Budish. 2011. The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes. *Journal of Political Economy* 119 (12 2011), 1061–1061. <https://doi.org/10.1086/664613>
- [14] IBM ILOG Cplex. 2009. V12. 1: User's Manual for CPLEX. *International Business Machines Corporation* 46, 53 (2009), 157.
- [15] Danny Cyril D Cruze. 2021. Amazon India announces creation of 1,10,000 new jobs ahead of festive season. <https://www.livemint.com/companies/news/amazon-india-announces-creation-of-1-10-000-new-jobs-ahead-of-festive-season-11632395844268.html>.
- [16] Delhivery. 2022. Delhivery warehouse hiring. <https://www.delhivery.com/services/ecommerce-logistics-services/>
- [17] Xiaotie Deng, Qi Qi, and Amin Saberi. 2012. Algorithmic Solutions for Envy-Free Cake Cutting. *Oper. Res.* 60, 6 (nov 2012), 1461–1476.
- [18] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness through Awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference* (Cambridge, Massachusetts) (ITCS '12). Association for Computing Machinery, New York, NY, USA, 214–226. <https://doi.org/10.1145/2090236.2090255>
- [19] EcomExpress. 2022. Be Our E-commerce Delivery Partners & Earn in Your Free Time. <https://ecomexpress.in/ship-with-us-campaign/>
- [20] Alejandro Escudero-Santana, Jesús Muñoz, Antonio Lorenzo Espejo, and María Luisa Muñoz. 2022. Improving E-Commerce Distribution through Last-Mile Logistics with Multiple Possibilities of Deliveries Based on Time and Location. *Journal of Theoretical and Applied Electronic Commerce Research* 17 (04 2022), 507–521. <https://doi.org/10.3390/jtaer17020027>
- [21] Indian Express. 2022. From Zepto to Blinkit, all the apps offering 10-minute delivery service in India. <https://indianexpress.com/article/technology/social/zomato-instant-zepto-blinkit-apps-that-offer-10-15-minute-delivery-service-in-india-7831106>.
- [22] Fairwork. 2021. Fairwork India Ratings 2020: Labour Standards in the Platform Economy. <https://bit.ly/3tsDlUa>.
- [23] Rajiv Gandhi, Samir Khuller, Srinivasan Parthasarathy, and Aravind Srinivasan. 2006. Dependent Rounding and Its Applications to Approximation Algorithms.
- [24] Corrado Gini. 1912. Variabilità e mutabilità.
- [25] Andrew V. Goldberg and Robert E. Tarjan. 1990. Finding Minimum-Cost Circulations by Successive Approximation. *Math. Oper. Res.* 15, 3 (1990), 430–466.
- [26] Nina Grgić-Hlača, Muhammad Bilal Zafar, Krishna P. Gummadi, and Adrian Weller. 2018. Beyond Distributive Fairness in Algorithmic Decision Making: Feature Selection for Procedurally Fair Learning. In *AAAI*. Article 7, 10 pages.
- [27] Anjali Gupta, Shreyans J Nagori, Abhijnan Chakraborty, Rohit Vaish, Sayan Ranu, Prajitt Prashant Nadkarni, Narendra Varma Dasararaju, and Muthusamy Chelliah. 2023. Towards Fair Allocation in Social Commerce Platforms. In *Proceedings of the ACM Web Conference 2023*. 3744–3754.
- [28] Anjali Gupta, Rahul Yadav, Ashish Nair, Abhijnan Chakraborty, Sayan Ranu, and Amitabha Bagchi. 2022. FairFoody: Bringing In Fairness in Food Delivery. In *AAAI*. 11900–11907.
- [29] Gurobi Optimization, LLC. 2022. Gurobi Optimizer Reference Manual. <https://www.gurobi.com>
- [30] Oliver P Hauser and Michael I Norton. 2017. (Mis) perceptions of inequality. *Current Opinion in Psychology* 18 (2017), 21–25.
- [31] Asheef Iqubbal. 2021. Food delivery workers in India: Emerging entrepreneurs or informal labour.
- [32] Manas Joshi, Arshdeep Singh, Sayan Ranu, Amitabha Bagchi, Priyank Karia, and Puneet Kala. 2022. FoodMatch: Batching and Matching for Food Delivery in Dynamic Road Networks. *ACM Trans. Spatial Algorithms Syst.* 8, 1, Article 6 (mar 2022), 25 pages. <https://doi.org/10.1145/3494530>
- [33] Ian Kash, Ariel D. Procaccia, and Nisarg Shah. 2013. No Agent Left behind: Dynamic Fair Division of Multiple Resources (AAMAS '13). International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 351–358.
- [34] Markus Knell and Helmut Stix. 2020. Perceptions of inequality. *European Journal of Political Economy* 65 (2020), 101927.
- [35] Alexander Kremer. 2021. Europe's quick-commerce startups are overhyped: Lessons from China. <https://techcrunch.com/2021/08/25/europes-quick-commerce-startups-are-overhyped-lessons-from-china>.
- [36] Bablu Kumar. 2020. Grievance. <https://twitter.com/BabluKu22667173/status/1257270087364509696>
- [37] R. J. Lipton, E. Markakis, E. Mossel, and A. Saberi. 2004. On Approximately Fair Allocations of Indivisible Goods (*EC '04*). Association for Computing Machinery, New York, NY, USA, 125–131. <https://doi.org/10.1145/988772.988792>
- [38] Ling Liu, Kumpeng Li, and Zhixue Liu. 2017. A capacitated vehicle routing problem with order available time in e-commerce industry. *Engineering Optimization* 49, 3 (2017), 449–465. <https://doi.org/10.1080/0305215X.2016.1188092>
- [39] Sepideh Mahabadi and Ali Vakilian. 2020. Individual Fairness for k-Clustering. In *Proceedings of the 37th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 119)*, Hal Daumé III and Aarti Singh (Eds.). PMLR, 6586–6596. <https://proceedings.mlr.press/v119/mahabadi20a.html>
- [40] Shruti Mehta. 2017. Zomato Restaurants Data. <https://www.kaggle.com/datasets/shrutimehta/zomato-restaurants-data>
- [41] Ritwik Moghe. 2022. Logistic. <https://bytes.swiggy.com/logistic-zones-for-assignment-48d9ce06c4a8>
- [42] Prasanna Mohanty. 2021. Gig economy: Good for companies, bad for workers. <https://www.fortuneindia.com/opinion/gig-economy-good-for-companies-bad-for-workers/106569>.
- [43] Nuno Mota, Negar Mohammadi, Palash Dey, Krishna P. Gummadi, and Abhijnan Chakraborty. 2021. Fair Partitioning of Public Resources: Redrawing District Boundary to Minimize Spatial Inequality in School Funding. In *Proceedings of the Web Conference 2021 (WWW '21)*. 646–657. <https://doi.org/10.1145/3442381.3450041>
- [44] Ashish Nair, Rahul Yadav, Anjali Gupta, Abhijnan Chakraborty, Sayan Ranu, and Amitabha Bagchi. 2022. Gigs with Guarantees: Achieving Fair Wage for Food Delivery Workers. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*. International Joint Conferences on Artificial Intelligence Organization, 5122–5128.
- [45] Govt. of Delhi. 2022. Enhancement of minimum wages. <https://labour.delhi.gov.in/content/order-nofno14202mwviiipart-file5972-5986-dated-14102022-enhancement-minimum-wages>
- [46] olist. 2019. Brazilian E-Commerce dataset. <https://www.kaggle.com/olistbr/brazilian-ecommerce>
- [47] John Olsson, Daniel Hellström, and Henrik Pålsson. 2019. Framework of Last Mile Logistics Research: A Systematic Review of the Literature. *Sustainability* 11 (12 2019), 7131. <https://doi.org/10.3390/su11247131>
- [48] Abraham Othman, Christos Papadimitriou, and Aviad Rubinfeld. 2014. The Complexity of Fairness through Equilibrium. arXiv:1312.6249 [cs.GT]
- [49] Shenle Pan, Vaggelis Giannikas, Yufei Han, Etta Grover-Silva, and Bin Qiao. 2017. Using Customer-related Data to Enhance E-grocery Home Delivery. *Industrial Management & Data Systems* 117 (09 2017), 1917–1933. <https://doi.org/10.1108/IMDS-10-2016-0432>

- [50] Gourab K Patro, Arpita Biswas, Niloy Ganguly, Krishna P Gummadi, and Abhijnan Chakraborty. 2020. Fairrec: Two-sided fairness for personalized recommendations in two-sided platforms. In *Proceedings of the web conference 2020*. 1194–1204.
- [51] Elisha A. Pazner and David Schmeidler. 1978. Egalitarian Equivalent Allocations: A New Concept of Economic Equity. *The Quarterly Journal of Economics* 92, 4 (1978), 671–687. <http://www.jstor.org/stable/1883182>
- [52] Pallavi Pengonda. 2021. Why covid-19 has been a shot in the arm for online food delivery firms. <https://www.livemint.com/market/mark-to-market/why-covid-19-has-been-a-shot-in-the-arm-for-online-food-delivery-firms-11611558506486.html>.
- [53] A Prakash, A Behera, B Sharma, N Relan, and V Kaul. 2022. Understanding Food Delivery Platform: Delivery Persons' Perspective.
- [54] Saritha Rai. 2021. Grocery Startup Founded by Teens Doubles Value to \$570 Million. <https://www.bloombergquint.com/business/grocery-startup-founded-by-teens-doubles-value-to-570-million>.
- [55] Ted Ralphs, L. Kopman, William Pulleyblank, and L.E. Trotter. 2003. On the Capacitated Vehicle Routing Problem. *Mathematical Programming* 94 (01 2003), 343–359. <https://doi.org/10.1007/s10107-002-0323-0>
- [56] Reuters. 2022. India's e-commerce firms hire more delivery workers for shopping season. <https://money.usnews.com/investing/news/articles/2022-08-23/india-e-commerce-firms-ramp-up-hiring-of-delivery-workers-for-shopping-season>.
- [57] Damián Reyes, Alan L. Erera, Martin W. P. Savelsbergh, Sagar Sahasrabudhe, and Ryan J. O'Neil. 2018. The Meal Delivery Routing Problem.
- [58] Priyanka Sahay, M. Sriram, and Sanghamitra Kar. 2021. Visit to a dark store: Here's what it takes to deliver that bag of groceries to you in 10 minutes! <https://www.moneycontrol.com/news/business/visit-to-a-dark-store-heres-what-it-takes-to-deliver-that-bag-of-groceries-to-you-in-10-minutes-7849131.html>.
- [59] David B. Shmoys, Éva Tardos, and Karen Aardal. 1997. Approximation Algorithms for Facility Location Problems (Extended Abstract). In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing (STOC '97)*. 265–274. <https://doi.org/10.1145/258533.258600>
- [60] Steni Simon. 2022. Zomato delivery agents strike for fair wages. <https://www.newindianexpress.com/cities/thiruvananthapuram/2022/mar/24/zomato-delivery-agents-strike-for-fair-wages-2433466.html>.
- [61] Walter Sinnott-Armstrong. 2022. Consequentialism. In *The Stanford Encyclopedia of Philosophy* (Winter 2022 ed.), Edward N. Zalta and Uri Nodelman (Eds.). Metaphysics Research Lab, Stanford University.
- [62] Vasudevan Sridharan. 2022. Workers on life in India's post-Covid, post-Great Resignation, gig economy: 'Like I'm working all the time'. <https://www.scmp.com/week-asia/economics/article/3163306/workers-life-indias-post-covid-post-great-resignation-gig>.
- [63] Mallika Srivastava and Madhur Raina. 2020. Consumers' usage and adoption of e-pharmacy in India. *International Journal of Pharmaceutical and Healthcare Marketing* 15, 2 (2020), 235–250.
- [64] Qian Sun, Welmer E. Molenmaker, Yikang Zhang, Yongfang Liu, and Eric van Dijk. 2022. Procedural fairness facilitates cooperative behavior by enhancing cooperative expectations. *Social and Personality Psychology Compass* 16, 12 (2022). <https://doi.org/10.1111/spc3.12711>
- [65] u/ChickenTeller. 2020. Grievance. https://www.reddit.com/r/doordash/comments/lsgw10/tony_please_explain_to_me_how_i_delivered_here_if/
- [66] Shai Vardi, Alexandros Psomas, and Eric Friedman. 2021. Dynamic Fair Resource Division. *Mathematics of Operations Research* 47 (09 2021). <https://doi.org/10.1287/moor.2021.1155>
- [67] Sabah Virani. 2021. Delivery workers prefer to work for smaller companies that offer better work conditions. <https://mumbai.citizenmatters.in/delivery-partners-the-biggest-apps-pay-the-least-26314>
- [68] Alicia Weber-Snyman and Johanna Badenhorst-Weiss. 2016. Challenges in last mile logistics of e-grocery retailers: A developing country perspective.
- [69] Barış Yıldız and Martin Savelsbergh. 2019. Provably High-Quality Solutions for the Meal Delivery Routing Problem. *Transportation Science* 53 (07 2019). <https://doi.org/10.1287/trsc.2018.0887>
- [70] Yuxiang Zeng, Yongxin Tong, and Lei Chen. 2019. Last-Mile Delivery Made Practical: An Efficient Route Planning Framework with Theoretical Guarantees. *Proc. VLDB Endow.* 13, 3 (nov 2019), 320–333. <https://doi.org/10.14778/3368289.3368297>
- [71] Sami Özarık, Lucas Veelenturf, Tom Van Woensel, and Gilbert Laporte. 2021. Optimizing e-commerce last-mile vehicle routing and scheduling under uncertain customer presence. *Transportation Research Part E: Logistics and Transportation Review* 148 (04 2021), 102263. <https://doi.org/10.1016/j.tre.2021.102263>