Matching Algorithms for Blood Donation

DUNCAN C MCELFRESH, University of Maryland, College Park
CHRISTIAN KROER, Facebook Core Data Science and Columbia University
SERGEY PUPYREV, Facebook Core Data Science
ERIC SODOMKA, Facebook Core Data Science
KARTHIK ABINAV SANKARARAMAN, Facebook
ZACK CHAUVIN, Facebook
NEIL DEXTER, Facebook
JOHN P DICKERSON, University of Maryland, College Park

Managing perishable inventory, such as blood stock awaiting use by patients in need, has been a topic of research for decades. Yet, most research focuses on the effective use of this scarce resource across the blood supply chain, and assumes the supply of blood itself can be impacted only via coarse policy levers. In this paper, empowered by the recently-deployed Facebook Blood Donation tool, we choose instead to model the first stage of the full blood supply chain—that is, the supply of blood itself—as a matching market. Here, potential blood donors are matched to donation centers by way of a central recommendation engine; that engine can at some cost prompt (e.g., via push notification) individual donors to donate to a preferred center or centers. Potential donors may have other constraints (e.g., maximum allowable frequency of prompts) or preferences (e.g., geographic, social) that should be taken into account. We develop policies for matching potential blood donors to donation centers under these constraints and preferences, and under simple models of demand shocks to the system. We test these policies in both computational simulations and real-world experiments on the Facebook Blood Donation tool. In both the simulated and real experiments, these matching-based notification policies significantly increase the expected number of blood donations.

1 INTRODUCTION

Blood is a scarce resource; its donation saves the lives of those in need. Countries approach blood donation in different ways, running the gamut from privately-run to state-run programs, with or without monetary compensation, and with varying degrees of public campaigns for action.[1] As such, blood donation rates differ across different countries; for example, approximately 3.2%, 1.5%, 0.8%, and 0.5% of the population donates in high-, upper-middle-, lower-middle-, and low-income countries, with varying rates of voluntary versus paid donors [37]. The case remains, though, that many patients do not have timely access to blood, especially in times of need. Thus, the World Health Organization (WHO) recommends that the blood supply chain—collection, testing, processing, storage, and distribution—be managed at a national level [37].

Optimization-based approaches to management of the blood supply chain have a rich history in the operations research and healthcare management literature. A recent paper [28] overviews over 100 publications in this space since the 1960s. The supply chain is roughly split into collection, testing & processing, storage & inventory, and distribution & transfusion [29]. Substantial research effort has gone into each of those segments [11, 24, 40]. Yet, we note that most optimization-based research in the initial *collection* stage of the blood supply chain has focused on *prediction* of blood supply (e.g., during a crisis). Given the ubiquity of social networks, in this work, we dovetail with that research by focusing instead on the *creation* of new blood supply via automated social prompts,

[1] Some examples follow. China maintains state control of its donation centers, which take a mix of captive-, quota-, and voluntary-based donations [20]. The US mixes state- and private-run donation that is primarily sourced via voluntary donations [28]. Brazil has seen a recent shift from remunerated to non-remunerated (aka voluntary) donation at its initially state-run, and now Federally-run, centers [6].

subject to the expressed preferences and constraints of potential donors and the overall donation system. We refer to this as the *donor recruitment* stage of the blood supply chain (see Figure 1).

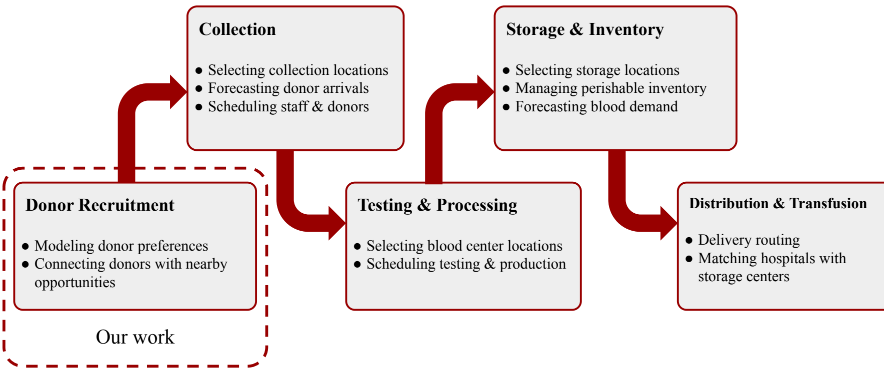


Fig. 1. Stages of the blood supply chain. Our work–donor recruitment–precedes the four stages of the blood supply chain as described in [28].

In this paper, we begin with an overview of related work on coordinating blood donors, and understanding their motivations (§2). Next we discuss the potential role that social media platforms might play in the blood donation process (§3); we present initial insights from the Facebook Blood Donation tool, and train a machine learning (ML) model for predicting when a notification will encourage donor action. Using these insights, we frame blood donor recruitment as a *matching* problem. Here, we draw on intuition from the literature investigating online bipartite matching under various arrival distribution assumptions, per-vertex and per-side budget constraints, and so on [18, 26]. Recent work in online matching has moved into more complex markets (e.g., crowdsourcing [21], rideshare [10], and kidney exchange [3, 4]), and has shown that—even in the face of theoretical intractability—sophisticated matching techniques can lead to improvements in economic efficiency. We aim to show a similar result here: we design a class of stochastic matching policies which allocate notifications *fairly* between recipients (§4).

This work is motivated by the recently developed Facebook Blood Donation tool[2], which connects millions of potential blood donors with opportunities to donate, in several countries around the world. We use anonymized data from the Facebook Blood Donation tool to build our matching model, and simulate various matching policies. Computational result indicate that even a relatively simple matching policy can greatly increase the magnitude of blood donations–while also treating recipients equitably. Finally, we run an online A/B test using the Facebook Blood Donation tool, to compare a (realistic) random baseline matching policy with a simple version of our proposed approach. We demonstrate that our simple policy substantially increases the expected number of donations, over the randomized baseline (§6).

2 RELATED WORK

In this paper we focus on the *recruitment* and *coordination* of potential blood donors, in order to meet ever-changing demand. While little work has focused on the coordination of blood donors and recipients, donor recruitment has been an active research area for decades; see [28] for a comprehensive review of prior work on the blood supply chain at-large. In §2.1 we discuss prior work related to donor recruitment; recently, in §2.2 we discuss similar work investigating the role

[2] `https://socialgood.fb.com/health/blood-donations/`

of the web and social media in blood donation. In §2.3 we briefly outline related approaches from the matching literature.

2.1 Understanding Donor Supply and Demand

Both extrinsic and intrinsic factors related to the donor and recipient play a role in understanding donation behavior. Extrinsic motivators such as influence from family or friends, or media and advertising, can increase donation rates [33]. In a controlled trial, Reich et al. [31] finds that empathetic messaging over phone and email can also encourage donation, more so than a material incentive. The medium used to recruit a donor is also important–whether in-person, over the phone, or in an email or letter [5]. Non-monetary incentives such as concert tickets or travel reimbursement can also influence donor behavior. In a review, Chell et al. [7] finds that these incentives are especially effective among young, first-time, and infrequent donors [7]; however some studies report the opposite effect among certain donor populations [15].

Intrinsic factors of the donor and perception of the donation experience also impact donation rates. Several studies report that a strong intention to donate and a sense of altruism are associated with higher donation rates [16, 33, 35]. Perceived difficulties in planning to donate can also impede donation rates [17, 35]. Many of these studies find that past donation behavior is a strong predictor of future donations [5, 16, 17] Both Schreiber et al. [32] and Yu et al. [38] find that first-time donors who donate frequently in the first year are more likely to be regular donors in the future.

A related area of research aims to predict supply and demand of donor blood, in order to avoid both shortages and waste. Techniques including forecasting [9], classification [27], and regression [14] have been used for decades to predict both supply and demand for donor blood.

Most importantly, many prior studies find that *different donors are motivated by different factors*. This suggests that personalized recruitment strategies–which respect diverse donor motivations, preferences, and perceived barriers to donation–will be more effective than a uniform recruitment strategy. As social influence and direct communication play an important role in donor recruitment, it is natural that social media (and generally, web-based apps) will play a role in the blood donation process. Indeed several recent studies have addressed the role of social media in blood donation; next we briefly outline this work.

2.2 Web Applications, Social Media, and Blood Donation

Web-based applications (apps) and social media platforms already play a substantial role in blood donor recruitment. Indeed, the American Red Cross, which claims to provide about 40% of transfused blood in the United States,[3] recently launched an app to connect blood donors with donation opportunities.[4] One review identified 169 different free mobile apps for blood donation [30]; though many of these apps have usability and privacy issues that may prevent widespread use. Regardless of whether these apps are available, it is important to understand whether donors will *use* these apps to find donation opportunities. Using a survey, Yuan et al. [39] finds that potential donors are receptive to using a web-based app to find and schedule blood donation appointments – and interest is especially high for young and infrequent donors. Respondents also expressed concern about privacy and receiving too many alerts from such an app; these are also primary concerns in our work.

Social media platforms also play an increasing role in blood donor recruitment. Sümnig et al. [34] finds that social media platforms (including Jodel and Facebook) are a major motivation for

[3]https://www.redcrossblood.org/donate-blood/how-to-donate/how-blood-donations-help/blood-needs-blood-supply.html
[4]https://www.redcrossblood.org/blood-donor-app.html

blood donors – especially for first-time donors. Similar studies have found that platforms including WhatsApp and Twitter influence donors in Saudi Arabia [2] and India [1]. A recent survey conducted with Facebook's partner blood banks found that 20% of donors said that Facebook influenced their decision to donate.[5]

2.3 Online Matching

The study of online bipartite matching problems goes back decades, to at least the seminal work of Karp et al. [23]. Mehta [26] provides an in-depth review of online matching models; we provide a short overview of related work here. One way to partition work in online matching is by *input model*—roughly, what kind of information the centralized matching mechanism is given about the way vertices arrive. Under the very conservative adversarial case (e.g., the model of [23]), the model is given no information about arrivals. Stochastic input models [13, 22] such as random order—where the input graph is chosen adversarially, but arrives randomly—and unknown IID—where vertices of particular types are drawn IID from an unknown distribution over types—are two less conservative input models. Our paper operates most closely to an even *less* conservative model—known IID, where vertices are drawn IID from a known distribution over types. This model can be appropriate when a matchmaker has access to vast amounts of data and can learn the distribution offline. This is common in the era of platform (rideshare, online labor) and advertising markets, and is true in our setting as well. Our setting is also related to another form of stochastic matching (see, e.g., [19]), where a matched edge may or may not exist (e.g., a push notification may or may not convert to an actual blood donation event). Finally, recent work (e.g., [10, 25]) looks at matching with *reusable* resources, primarily motivated by rideshare where a driver is matched for the course of a trip, and then reappears. In our setting, blood donors are matched and then taken offline for some number of time periods before reappearing as available, potential matches.

3 FACEBOOK'S ROLE IN BLOOD DONATION

The advent of global social networks offers a unique opportunity to recruit and coordinate massive numbers of donors, in order to meet a large and unpredictable demand for donor blood. The *Facebook Blood Donation Tool* aims to seize this opportunity – using the widespread use of its online platform to connect blood donors with nearby recipients (see Figure 2). Donors can also opt-in to receive *notifications* about nearby donation opportunities. This tool is available in several countries around the world, including Bangladesh, Brazil, India, Pakistan, and the United States; more than 35 million Facebook users have registered with this tool.[6]

In this paper we focus on a small but important feature of the Blood Donation tool: automatic donor notifications. Our primary goal is to *increase the number of blood donations around the world* by carefully selecting *which opportunity* to notify each donor about, and *when* to notify them. We frame this question of donor notifications as an *online matching problem*, which we formalize in §4. One might ask whether such a complicated approach is warranted in this setting: perhaps it does not matter how and when donors are notified. To address this question, we first ask: how can tell whether a Facebook user donates blood after we notify them?

3.1 Measuring Donation: Meaningful Action.

To design notifications that effectively encourage blood donation, it is necessary to observe *when* a donor donates. However social networking platforms like Facebook cannot directly observe a user's action outside the platform. As a proxy, Facebook can instead observe when a donor takes

[5]Facebook Newsroom article: `https://about.fb.com/news/2019/06/us-blood-donations/`.

[6]`https://about.fb.com/news/2019/06/us-blood-donations/`.

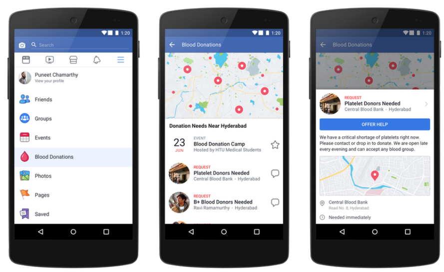


Fig. 2. The Facebook Blood Donation tool interface, where users can search for donation opportunities, and opt-in to receive notifications about nearby opportunities as they arise. (Source: `https://about.fb.com/news/2018/06/making-it-easier-to-donate-blood`.)

meaningful action toward donation–after being notified. In our context, meaningful actions include actions like creating a reminder to donate, or calling a blood bank (these actions are only observed if taken within the Facebook platform).

Meaningful action (MA) has been validated as a reasonable proxy for actual blood donations, using surveys of Facebook's partner blood banks and hospitals. In the remainder of this paper, we focus on increasing the number of donor MAs as a proxy for increasing the number of donations. Our goal is to design a notification policy that chooses both (a) *when* to notify a donor, and (b) *which donation opportunity* to notify them about. The next step in designing this policy is to understand *which* notifications are likely to prompt donor MA. We begin with some high-level observations.

3.2 What Notifications Encourage Meaningful Action?

As an initial analysis we consider hundreds of millions of notifications sent to donors using Facebook Blood Donation tool, over a one-month period. Below we describe some high-level observations of this notification data; we leave a deeper analysis to future work:

(1) **Meaningful action is rare:** between 3% and 4% of all notifications lead to meaningful action.
(2) **More-engaged donors are more likely to take meaningful action:** Donors who tend to use Facebook every day are about 43% more likely to take action than those who use Facebook about once per week.
(3) **New users are more likely to take action:** donors who joined Facebook within the last year are about 35% more likely to take action that those who have been users for longer.
(4) **Older donors are more likely to take action:** donors over 30 years old are about 22% more likely to take action than donors under 30.
(5) **Donors are more likely to take action if they are notified about a nearby opportunity:** Donors who are notified about opportunities who are less than 3km away are 20% more likely to take action than those who are notified about further-away opportunities.
(6) **Donors are more likely to take action if they haven't been notified recently:** Donors who haven't been notified about a donation opportunity in the past 60 days are about 12% more likely to take action than those who have been notified in the past 60 days.

We emphasize that observations have been reflected in prior studies. (2) reflects the observation of [33] and Sümnig et al. [34] that social pressure and influence from family or friends can increase donation rates. (5) reflects the finding of Van Dongen et al. [35] and Godin et al. [17] that logistical barriers to donation can impede donation rates. (6) reflects the finding of Yuan et al. [39] that blood donors are concerned about receiving too many notifications.

The likelihood of donor MA varies significantly, and is correlated with several features of both the blood donor (e.g., when they were last notified) and donation opportunity (e.g., location). To better capture these dependencies, we train a predictive model for estimating likelihood of donor MA, using all available data from prior notifications. This model will be used in both our offline and online experiments.

3.3 Machine Learning Model for Donor Action

To develop a machine learning (ML) model of donor action, we use all prior notifications sent by the Facebook Blood Donation tool. This model takes an individual notification as input, and predicts the *probability* that the donor will take action. Each notification is represented by a set of *features* of both the donor and the donation opportunity (i.e., the inputs to the model); before being deployed, this ML model and application passed through an internal review to protect user privacy.

Prior to training this model, we use industry-standard feature selection techniques to identify the most important features for predicting donor MA; these features are (in decreasing order of importance, with importance percentage in parenthesis): (1) whether the donor recently took meaningful action (18%), (2) donor age (8.5%), (3) donor city (7.5%), (4) the number of Facebook friends the donor has (7.3%), (5) the distance between donor and recipient (6.8%). Other relevant features include the number of local donors (6.5%), number of times a donor has viewed the hub in the last 30 days, and the number of days since the donor's last notification.

Using the selected features, we train a gradient boosted decision tree (GBDT) model. We use standard parameter-sweep techniques to obtain the learning rate of 0.1, 120 trees, a maximum tree depth of 5 and a maximum number of leaves of 120. This model is trained using 10-fold cross-validation on 80% of the the training data and an additional 10% for validation; it achieves an AUC of 0.66 and logistic loss of 0.45, averaged over all training folds. Training this model is particularly challenging because of the small number of "positive" examples (i.e., the number of donor MAs). Figure 3 shows a histogram of prediction scores for all training data. Most prediction scores are between 0-10%, with an average of 3.43%–which closely the observed likelihood of MA.

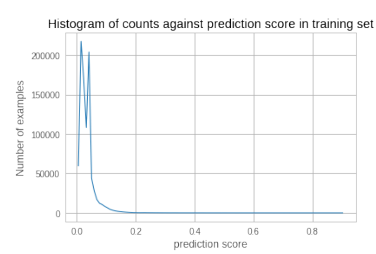


Fig. 3. Distribution of labels in the training dataset

We use this model to estimate *how likely* it is that a donor will take action, when notified about a particular donation opportunity. Next we describe how to use this model to design a notification policy: by framing blood donor recruitment as a *matching* problem.

4 BLOOD DONATION NOTIFICATION AS MATCHING

In this section we formalize the the donation scenario as a *matching problem*. In §4.1 we outline the matching framework, in §4.2 we describe two objectives for this problem, and in §4.3 we show that this problem is NP-complete under most conditions.

4.1 Matching Framework

We represent a blood donation problem as a weighted bipartite *donation graph* $G = (U, V, E)$, with donors $u \in U$ and donation opportunities (or *recipients*) $v \in V$. [7] Each vertex has a set of *attributes*

[7] We use the terms "donors" and "recipients" as shorthand for *prospective* donors and recipients. Facebook does not make any determination about a person's eligibility to donate blood; these are potential donors who sign up to receive notifications of blood donation opportunities.

(e.g., blood type, geographical location, and so on); donor and recipient attributes determine whether a donor u can donate to a recipient v (i.e., whether u and v are *compatible*). Compatible pairs (u, v) are connected by edges $e = (u, v) \in E$; we denote all edges adjacent to vertices $u \in U$ ($v \in V$) as $E_{u:}$ ($E_{:v}$). If an edge $e = (u, v)$ exists, then donor u can be *notified* about v.

Time Dependence: The donation graph G is inherently *dynamic*, in that some edges may only be in G temporarily. [8] We discretize time into days $t \in \mathcal{T} \equiv \{1, \dots T\}$, with a finite-time horizon T. Let $V(t)$ and $E(t)$ be the set of available recipients and edges at time t; $E_{u:}(t)$ ($E_{:v}(t)$) are defined similarly as the edges adjacent to donor u (recipient v). Furthermore, we assume that changes to the edge set $E(t)$ are due only to dynamic *demand*; that is, we assume that donors are always available to receive donations. We assume that recipient arrivals (and edge sets) are unknown for all future time steps. However we assume a known *distribution* of arrivals defined by p_{vt}, the probability that recipient v is present at time t. For ease of exposition, we set $p_{vt'}$ to 0 (1) for all recipients that were absent (present) on *observed* (i.e., past and present) time steps t'.

Edge Weights: Each edge $e = (u, v)$ has weight equal to the *probability* that donor u donates to recipient v once notified (i.e., the predicted MA likelihood, see §3.3). Edge weights may change with time, *and* may depend on the donor's notification history. For these reasons we index edge weights w_{et} by both edge e and time t.

Recipients: We consider both *static* recipients $S \subseteq V$, such as blood banks and hospitals, and *dynamic* recipients $D \subseteq V$, such as blood drives or emergency requests. Static recipients are available during *all* time steps, and edges into these recipients are always available (i.e., $E_{:s}(t) = E_{:s}$). In contrast, events arrive in an online manner, and are available only during a fixed subset of time steps, denoted by $T(d) \subseteq \mathcal{T}$. For time steps outside of this duration (any $t' \notin T(d)$) the edge set is empty – i.e., $E_{:d}(t') \equiv \emptyset$. In many cases it may be advantageous to "plan ahead", and reserve some donors for potential future events. This is not simply a modeling choice: unpredictable natural disasters and health emergencies can cause spikes in the demand for donor blood that require additional supply.

Donors: Once a donor signs up with the Facebook Blood Donation Tool, we say they are *available* to receive notifications (i.e., to be matched) once every K days (for some reasonable K, around two weeks). This limitation is motivated by respect for donors, and also the hard constraint on the frequency of blood donations allowed by the country or municipality.[9] Furthermore, initial analysis of past notifications suggests that it is beneficial to notify donors periodically; we summarize these observations with the following assumptions.

Assumption 1. *If a donor is notified at time t, then $w_{et'} = 0$ for $t' = t + 1, \dots, t + K$.*

In other words, donors are very unlikely to donate if we notify them more than once every K days; in this paper we use $K = 14$.

Decision Variables: At each time t, we decide which notifications to send (i.e., which *edges*). Let $x_{et} = 1$ if we decide to notify a donor u using edge $e = (u, v)$, and $x_{et} = 0$ otherwise.

4.2 Matching Objectives

In this initial work we propose two matching objectives, to represent (a) the overall number of donations, and (b) the equitable treatment of recipients.

[8]In this initial work, we assume the set of potential donors and donation centers do not change, although this *longer-term dynamism* is certainly interesting to consider as future research.

[9]Typically 8 weeks or longer; see https://www.redcrossblood.org/faq.html.

Efficiency Objective: Overall Number of Donations. Our first objective is to maximize the expected number of total blood donations (henceforth *efficiency*). Using the notation from above, we can express the overall efficiency of a matching policy as

$$\sum_{t \in [T]} \sum_{e \in E(t)} w_{et} x_{et}.$$

In general, we aim to select notifications (using x_{et}) to maximize this quantity.

Fairness: Proportional Allocation. We also want to treat recipients *equitably*. We take a utilitarian approach to express fairness in this setting: let $U_v \equiv \sum_{t \in \mathcal{T}} \sum_{e \in E_{:v}(t)} w_{et} x_{et}$ be the *utility* awarded to recipient v at time t (i.e., the total edge weight of all notifications about recipient v – the *expected number of donations* to recipient v). In general we want recipients to receive an equal share of donations. However not all recipients can receive the same number of donations. Recipients may have different numbers of compatible donors (e.g., due to their location), or different edge *weights* (e.g., due to donor preferences or recipient accessibility). For this reason we consider the *scaled utility* for each donor, $U_v^s \equiv U_v / m_v$, where m_t is a normalization score defined for each recipient v. In our model, n *equitable* notification policy should award roughly the same scaled utility U_v^s to all recipients v over the entire time horizon $\mathcal{T}$.

Normalization: Uniform Random Allocation as the "Equitable" Outcome. The fairness objective defined above depends strongly on the choice of *normalization*. We consider a normalization score motivated by the concept of proportional fair division [12]. In this setting, the "perfectly equitable" matching is one in which each donor is matched uniformly at random to a recipient, as soon as the donor is available (K days have passed since the donors' last notification). We define m_v as the expected weight matched to recipient v under this uniform random matching. Calculating m_v exactly would require extensive sampling or simulation, because edge weights depend on both time and prior match decisions. Instead, at each time step t', we approximate the uniform-random matching by only considering the next K time steps; due to Assumption 1, each donor is matched at most once in this period, and thus m_v can be calculated exactly. Let n_u be the number of edges adjacent to donor u during this time period:

$$n_u \equiv \sum_{t=t'}^{t'+K} |E_{u:}(t)|.$$

We calculate m_v as follows:

$$m_v \equiv \sum_{t=t'}^{t'+K} \sum_{e=(u,v) \in E_{:v}(t)} \frac{w_{et}}{n_u}$$

In practice this notion of proportional may be too strict–and may decrease the overall welfare of the matching market. For this reason use the following notion of *relaxed proportional fairness.*

Definition 4.1 (γ-Proportionally Fair Allocation). Let U_v^s be the scaled utility awarded to recipient v. This assignment is γ-proportionally fair if the following inequalities hold for all recipients $v \in V$:

$$\frac{\gamma}{|V|} \sum_{v' \in V} U_{v'}^s \leq U_v^s \leq \frac{1/\gamma}{|V|} \sum_{v' \in V} U_{v'}^s$$

for some $\gamma \in (0, 1]$. For assignments where there is no $\gamma \in (0, 1]$ for which these constraints hold, we say $\gamma = 0$.

That is, the assignment is γ-proportionally fair if U_v^s is within fraction γ and $1/\gamma$ of the mean scaled utility over all recipients. Setting $\gamma = 1$ requires that all recipients receive the same scaled utility, while setting $\gamma = 0$ makes these constraints non-binding (because $U_v^s \geq 0$). Throughout this paper we will also refer to a recipient's *proportional outcome* α_v, defined as

$$\alpha_v \equiv \frac{U_v^s}{\frac{1}{|V|} \sum_{v' \in V} U_{v'}^s}.$$

Note that the proportional fairness constraints are equivalently expressible in terms of proportional outcome, as

$$\gamma \leq \alpha_v \leq 1/\gamma \quad \forall v \in V.$$

Before discussing solution approaches to this matching problem, we first show that in most cases this matching problem is NP-hard.

4.3 NP Hardness

The following theorem states that the matching problem described in § 4 is strongly NP-hard when fairness is considered, even for a single time step.

Theorem 4.2. *The problem of maximizing overall donations subject to γ-proportionally-fair allocation and integer-only allocation is strongly NP-hard for any fixed $\gamma \in (0, 1]$, even when $T = 1$.*

The proof of Theorem 4.2 is given in Appendix D. While this matching problem is hard in general, we never intend to solve it directly. The huge number of blood donors and recipients would likely make this problem intractable even with an efficient linear programming formulation. Instead we study several matching *policies* which are designed to be tractable in real settings, while using the same objective and constraints described in this section.

5 MATCHING POLICIES FOR BLOOD DONATION

In this section we propose a class of stochastic matching policies to match blood donors donors and recipients. In §5.1 we propose a stochastic matching policy which use distributional assumptions of future demand; in §5.2 we propose a *myopic* variant of this policy, which ignores future demand; in §5.3 we propose a final variant which ignores both future demand and fairness constraints.

5.1 LPMatch(γ): A Stochastic Policy with a Limited Planning Horizon

Next we propose LPMatch(γ), a realistic stochastic policy based on the LP relaxation of an offline-optimal problem. At each time step $t \in \mathcal{T}$, this policy considers a *limited planning horizon* including the next $H < K$ time step, denoted by $T^H \equiv \{t', \dots, t' + H\}$. In addition, this policy includes γ-proportional fairness constraints (see § 4) for *only* the planning horizon T^H. In our setting, future demand is unknown (i.e., the set of available recipients for future time steps). However we assume distributional knowledge of future arrivals (see § 4); we refer to a fixed sequence of recipient arrivals as a demand *realization*. For any particular demand realization we might formulate an *offline-optimal* matching policy – i.e., a mixed-integer linear program (MILP) which "perfectly" matches donors with recipients. In this section we formalize an LP relaxation of the general offline-optimal MILP, which upper bounds the objective value for *any* possible realization. LPMatch(γ) uses an optimal solution of this LP to derive a stochastic matching policy – similar to the approach of Dickerson et al. [10] for matching passengers with drivers in a rideshare setting.

At each time step t, LPMatch(γ) re-solves this LP relaxation to construct a matching for the current time step. Note that using a longer time horizon introduces several complications: with

$H > K$, donors may be matched more than once during T^H, and weights can depend on decision variables x_{et} (see Assumption 1).

Next we formalize an LP relaxation for the offline-optimal MILP using the following parameters:

- $p_{vt} \in [0, 1]$, the probability that recipient v is present at time t. These probabilities may depend on v's history (e.g., whether or not v was present during the previous time step).
- Edge weights $w_{et} \geq 0$ – i.e., the probability that donor u takes action when notified about recipient v along edge $e = (u, v)$.
- $\gamma \in [0, 1]$, the level of proportional fairness required of the outcome.

Decision variables $x_{et} \in \mathbb{R}_+$ are defined for all edges $e \in E(t)$ and all time steps $t \in \mathcal{T}$. In the offline-optimal MILP, $x_{et} \in \{0, 1\}$ is 1 if e is matched at time t and 0 otherwise; in the LP formulation, x_{et} is the *probability* that e is matched at time t, for any demand realization (i.e., in any offline-optimal MILP).

Constraints. The following constraints require that each donor is matched *at most* once during the time period T^H:

$$\sum_{t \in T^H} \sum_{e \in E_{u:}(t)} x_{et} \leq 1 \quad \forall u \in U.$$

In any offline-optimal matching, a recipient v can only be matched if the recipient is present at time t; the following constraint upper-bounds the probability that recipient v is matched at time t with p_{vt} – the probability that v is present at t:

$$x_{et} \leq p_{vt} \quad \forall e \in E_{:v}(t),\ v \in V,\ t \in T^H.$$

The above constraints ensure that the optimal objective to the LP upper-bounds the offline-optimal objective, for any realization of stochastic online demand. Finally, the following constraints enforce relaxed proportional fairness over fixed time horizon T^H. Let V_H denote the recipients present at any time during T^H:

$$\frac{\gamma}{|V_H|} \sum_{v' \in V_H} U^s_{v'} \leq U^s_v \leq \frac{1/\gamma}{|V_H|} \sum_{v' \in V_H} U^s_{v'}$$

with auxiliary variables

$$U^s_v = \frac{1}{m_v} \sum_{t \in T^H} \sum_{e \in E_{:v}} x_{et} w_{et},$$

where m_v is calculated using a uniform random matching over time horizon T^H, as described in Section 4. Note that these constraints consider only a subset T^H of the complete timeline $\mathcal{T}$. For this reason, the notion of fairness in `LPMatch`(γ) is localized in time (to rolling horizon T^H); the matching result from `LPMatch`(γ) is not expected to obey the "gloabl" fairness constraints described in § 4.2.

Problem 3 gives the complete LP relaxation.

$$
\begin{aligned}
\max \quad & \sum_{t \in T^H} \sum_{e \in E(t)} w_{et} x_{et} & \\
\text{s.t.} \quad & x_{et} \in [0,1] & \forall e \in E(t)\ t \in T^H \\
& U^s_{vt} \in \mathbb{R} & \forall v \in V(t), t \in T^H \\
& x_{et} \leq p_{vt} & \forall v \in V(t), e \in E_{:v}(t), t \in T^H \\
& \sum_{t'=t}^{t+k} \sum_{e \in E_{u:}(t)} x_{et} \leq 1 & \forall u \in U, t = \{1, \dots, T^H - k\} \\
& U^s_v = \frac{1}{m_v} \sum_{t \in T^H} \sum_{e \in E_{:v}(t)} x_{et} w_{et} & \forall v \in V_H \\
& U^s_v \leq \frac{1/\gamma}{|V_H|} \sum_{v' \in V_H} U^s_{v'} & \forall v \in V_H \\
& U^s_v \geq \frac{\gamma}{|V_H|} \sum_{v' \in V_H} U^s_{v'} & \forall v \in V_H
\end{aligned}
\tag{1}
$$

Algorithm 1 describes the matching policy LPMatch(γ). This algorithm takes as input p_{et} and w_{et}, both of which may be updated between time steps. After solving Problem 3, LPMatch(γ) creates a matching distribution for each recipient, with probabilities proportional to the optimal LP solution x^*_{et}. With probability $\left(1 - \sum_{e \in E_{u:}(t')} x_{et}\right)$ the recipient is left unmatched.

ALGORITHM 1: LPMatch(γ)

---

input : t' (current time step), U (available donors), $V(t')$ (available recipients), $p_{vt}\ \forall\ v \in V_H,\ t \in T^H$ (recipient arrival probabilities); $w_{et}\ \forall\ e \in E(t), t \in T^H$ (edge weights)
output: E^* (matched edges for time t)

$x^*_{vt} \leftarrow$ optimal solution to Problem 3
$q_v \leftarrow \sum_{e \in E_{u:}(t')} x_{et'}$ `// total probability of matching` v

for $u \in U$ **do**

$P[e'] \leftarrow \begin{cases} x_{et^*}/q_v, & e' = e, \forall e \in E_{u:}(t^*) \\ 1 - q_v, & e' = \emptyset \end{cases}$ `// matching distribution for recipient` u

sample $e^* \sim P$

end
return E^*

Next we consider two variants of LPMatch(γ): a *myopic* policy LPMyopic(γ) which consider only the current time step, and a myopic *max-weight* policy MaxWeight which ignores fairness constraints.

5.2 LPMyopic(γ): Myopic Stochastic Matching

There are some cases where it is necessary or preferable to ignore future demand. For example, if the number of donors or recipients is too large, Problem 3 may not be solvable in a reasonable amount of time. If parameters p_{vt} or w_{et} are unknown or very uncertain, it may be unwise to use these parameters to guide a matching policy. In this case we propose a variant of LPMatch(γ), referred to as LPMyopic(γ), which considers *only* the current time step, t'. LPMyopic(γ) is equivalent to LPMatch(γ) in that is uses Algorithm 1, but considers *no* time horizon – i.e., $T^H \equiv \{t'\}$.

We emphasize two key differences between LPMatch(γ) and LPMyopic(γ). First, since LPMyopic(γ) considers only the current time step, all demand is assumed to be known (i.e., $p_{vt'} \in \{0, 1\}$). Second, the fairness constraints in Problem 3 ignore all future demand, and are thus far more "localized" in time than those of LPMatch(γ).

Next we discuss another modification of LPMatch(γ) which does not consider fairness.

5.3 MaxWeight: Max-Weight Stochastic Matching Policy

The final policy we consider is MaxWeight – a myopic policy which ignores fairness constraints. Removing these fairness constraints is equivalent to *completely* relaxing the γ-proportional fairness constraints in Problem 3 Similar to LPMatch(γ) and LPMyopic(γ), MaxWeight uses Algorithm 1 to match donors and recipients, however rather than solving Problem 3, MaxWeight uses the following LP relaxation, Problem 2

$$
\begin{aligned}
\max \quad & \sum_{t \in T^H} \sum_{e \in E(t)} w_{et} x_{et} \\
\text{s.t.} \quad & x_{et} \in [0, 1] && \forall e \in E(t)\ t \in T^H \\
& x_{et} \leq p_{vt} && \forall v \in V(t), e \in E_{:v}(t), t \in T^H \\
& \sum_{t \in T^H} \sum_{e \in E_{u:}(t)} x_{et} \leq 1 && \forall u \in U
\end{aligned}
\tag{2}
$$

Note that under certain circumstances, this LP is expressible as a maximum network flow problem. In particular, if all demand is known (i.e., $p_{vt} \in \{0, 1\}$, as is the case with $H = 0$), then this LP is equivalent to network flow, and has an integral solution. We characterize this case in Appendix C.

Next we describe the data and models used in both our offline and online experiments.

6 OFFLINE SIMULATIONS

In these experiments we compare the performance of three realistic matching policies in a simulated computational setting. The policies we test in this section is *realistic*, in the sense that they only have access the *current* state of the matching problem–and not about the future. These policies notify each donor exactly once every $K = 14$ days (i.e., as soon as the donor becomes "available"). All policies have access to the same set eligible edges (where the donor is less than 15km from the recipient), provided by the edge weight model described in Section 3.3.

That is, the only difference between these policies is in *which recipient* they choose to notify each donor about; we compare three different policies: MaxWeight (myopic max-weight matching), Rand (each donor is randomly notified along an available edge), LPMyopic(γ) (the myopic LP matching policy with γ-relaxed fairness constraints).

We test each of these policies on several different cities, each with hundreds of thousands of donors, and each with around 100 recipients. In these experiments we only consider *offline* recipients (e.g., hospitals and large blood banks) which are always available. To emulate a real notification setting where donors become available on a rolling basis, we randomly assign each donor a "start day" between the 1^{st} and the K^{th} day of the simulation. We run these simulations for a 15-day period, using data from November 2019.

6.1 Metrics: Efficiency and Fairness

There are two outcomes of each policy we wish to highlight: *efficiency* (i.e., total edge weight, the total number of expected MAs), and *fairness* (i.e., the distribution of recipients' proportional outcomes α_v). Let x^*_{et} be the final decision variables for a particular policy (i.e, x^*_{et} is 1 if edge e is matched at time t, and 0 otherwise). We calculate two sets of metrics for each policy, using the final decision variables x^*_{et}: first, the total edge weight W (i.e., the total number of expected donations):

$$W \equiv \sum_{t \in \mathcal{T}} \sum_{e \in E(t)} w_{et} x^*_{et}.$$

Second, the proportional outcome for each agent α_v, defined as

$$\alpha_v \equiv \frac{U^s_v}{\frac{1}{|V|} \sum\limits_{v' \in V} U^s_{v'}},$$

with

$$U_v \equiv \frac{1}{m_v} \sum_{t \in \mathcal{T}} \sum_{e \in E_{:v}(t)} w_{et} x^*_{et}.$$

Note that notification policies Rand and LPMyopic(γ) are stochastic–their outcomes may be slightly different depending on a random seed set by the optimizer. For this reason, we simulate 20 different *realizations* of Rand and LPMyopic(γ); we calculate the *average* W and α_v for each recipient, over all 20 realizations.

Visualizing Policy Outcomes: Pareto Frontier. To visualize how these policies trade-off between total edge weight (the overall number of donations) and fairness to recipients, we plot each simulation outcome as a point, where the vertical axis corresponds to total matched weight, and the horizontal coordinate corresponds to the median recipient scaled utility α_v. (For the stochastic policies, this is the median of the *average* α_v for each recipient, over all 20 realizations.) Figure 4 shows the outcomes for each policy, for 12 different cities around the world.

As expected, MaxWeight always maximizes total edge weight, but often achieves a very low median α_v. In fact, in several of the cities (1, 2, 3, 5, 8, 10, 11, 12) MaxWeight results in a median α_v of zero. In these cases, *at least half half* of all recipients were left unmatched – i.e., no donors were notified about these recipients.

On the other hand, Rand tends to maximize α_v (often achieving a median value around 1 – or "perfect" proportional fairness), but with a significantly lower total matched weight.

Somewhat surprisingly, our proposed policy LPMyopic(γ) often achieves a *Pareto improvement* over both Rand and MaxWeight. That is, there is often a γ such that either

- LPMyopic(γ) achieves a *greater* total weight than Rand, without a lower median α_v (nearly all Cities in Figure 4), or
- LPMyopic(γ) achieves a *greater* median α_v than MaxWeight, without a lower W (see Cities 1, 3, 6, 8, 11, 12).

Before discussing our online experiments, we highlight some lessons learned from these simulations, which motivate our online experiments. First, maximizing edge weight (i.e., policy MaxWeight) should increase the number of MAs by 10-15% over random notification, according to our predictive model. This comes at a cost to fairness – in several instances, MaxWeight leaves at least half of all recipients unmatched. Random notification (Rand) is often fair to recipients, but with much lower matched weight. Our randomized fair policy (LPMyopic(γ)) moderates between the MaxWeight and Rand policies, and often achieves a Pareto improvement over both MaxWeight and Rand. The trade-off between efficiency and fairness is controlled smoothly using parameter (γ).

7 ONLINE EXPERIMENTS

As a proof-of-concept, we compare the max-weight matching policy (MaxWeight) to the random baseline policy (Rand, which is similar in behavior to the notification policy currently used by the Blood Donation system), in an online experiment. The goal of this experiment is to answer the question: *can we increase the overall number of donor meaningful actions* by carefully selecting *which*

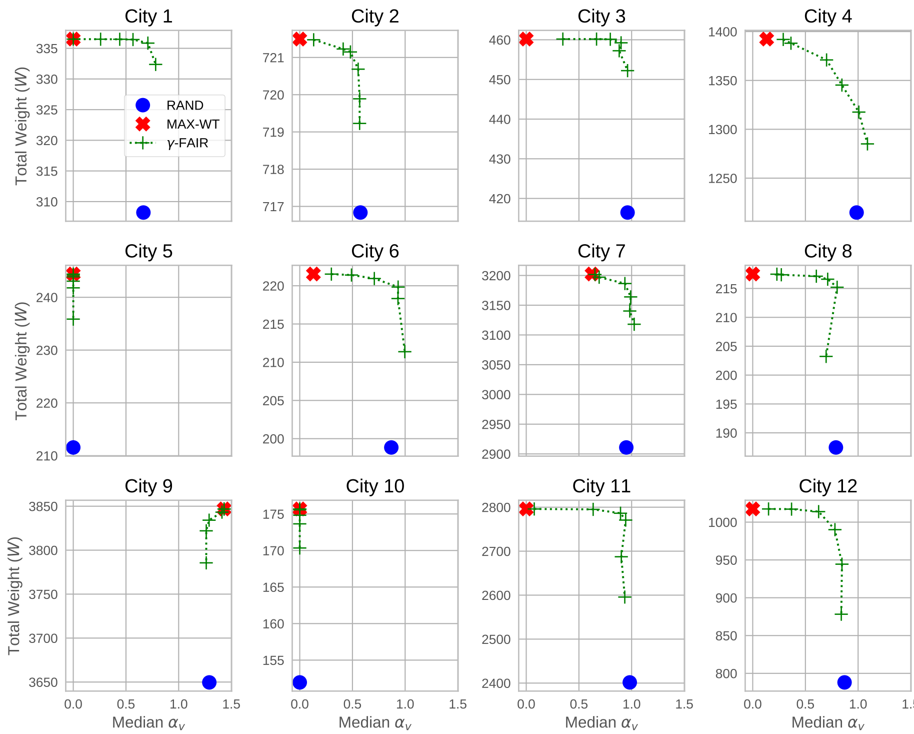


Fig. 4. Simulation outcomes for 12 cities around the world, each with hundreds of thousands of donors and around 100 recipients. Each plots shows the total matched weight (W) and median proportional recipient outcome α_v for notification policy Rand (blue circles), `MaxWeight` (shown as red "X"), and `LPMyopic`(γ) (shown as green "+") for $\gamma \in \{0.1, 0.3, 0.5, 0.9, 1.0\}$. Note that `LPMyopic`($\gamma$) effectively moderates between a `MaxWeight` and a the "fair" policy Rand, while consistently matching a much higher edge weight than Rand.

recipient to notify each donor about. Both of these policies notify donors once every $K = 14$ days; they only differ in *how* they select a recipient to notify each donor about. Rand selects a nearby recipient at random, while `MaxWeight` selects a nearby recipient with the greatest likelihood of donor MA–according to our predictive edge weight model.

To compare these policies we design a randomized an online experiment, including hundreds of thousands of donors registered with the Facebook Blood Donation tool. We randomly partition these donors into a control group (who were notified using policy Rand) and a test group (who were notified using policy `MaxWeight`). As in our simulations, we include *only* static recipients (e.g., hospitals and large blood banks), who are always available to receive donations.

Potential Impact on Donors and Recipients. This experiment was approved by an internal review board. We emphasize that the impact of these experiments is minimal: the only difference between the test and control group in this experiment is *which* donation opportunity the donor is notified about. The impact on blood recipients is less clear: due to our experimental design we cannot effectively measure the *fairness* of each notification policy in a meaningful way. However it is possible that any optimization-based matching policy (e.g., `MaxWeight` or `LPMyopic`(γ)) prioritizes certain recipients over others. This may marginalize recipients in rural areas or those with a limited

Table 1. Online Experiments - Number of notifications (#Notifs) and meaningful actions (#MA), over the online experiment. Notifications are separated into those sent to donors with only one compatible recipient (1R), and those sent to donors with two or more compatible recipients (≥2R). Wilson score intervals are for the percentage of notifications that lead to MA are presented as $C \pm R/2$, where the 95% confidence interval is $[C - R/2, C + R/2]$.

	Notif.	Control (Rand)			Test (MaxWeight)		
	Group	#MA	#Notifs	%MA	#MA	#Notifs	%MA
Experiment	1R	10,534	215,544	4.7 ± 0.1	10,755	214,841	4.8 ± 0.1
	≥2R	17,479	469,887	3.6 ± 0.1	18,030	459,708	3.8 ± 0.1

Facebook presence. More thorough analysis of these impacts is necessary before more widespread adoption of these policies.

7.1 Online Experiment Results

This experiment ran from Nov. 23 to Dec. 17, 2019 (26 days); in total, 1,359,980 donors were notified from both the test and control group. In this experiment many donors had only one compatible recipient – in this case, the donor was *always* notified about this recipient, regardless of the notification policy. For clarity, we distinguish between notifications sent to donors who had only one compatible recipient (1R), and those sent to donors with two or more compatible recipients (≥2R). Thus we only expect to observe a difference between control and test groups for ≥ 2R notifications; we expect the same outcome for (1R) notifications. Table 1 shows the number of notifications and meaningful actions for notifications of each type (1R and ≥ 2R), in both the test and control group.[10] (Note that only ≥ 2R notifications are relevant for comparing the test and control groups, though we report both for transparency.) The key result in these tables is the percentage of notifications that led to meaningful action (%MA, a number on $[0, 100]$). Due to the small sample size for certain cities we report the Wilson score interval for %MA as $C \pm R/2$, where $[C - R/2, C + R/2]$ is the 95% confidence interval.

In the remaining discussion we consider only the ≥ 2R notifications, as there is no difference between the test and control group for 1R notifications. For the overall experiment, %MA is about 5% higher for `MaxWeight` than for `Rand`. However for individual cities, the number of notifications and MAs is too small to draw conclusions. To better understand the differences between the control and test groups, we use two statistical tests to compare the *overall* results.

Overall Comparison. We use both a two-sided and one-sided Chi-square test to compare %MA (≥ 2R notifications only) for the control and test groups, over all notifications sent during this experiment. Let $P_{\texttt{Rand}}$ and $P_{\texttt{MaxWeight}}$ represent %MA for the control (`Rand`) and test (`MaxWeight`) groups, respectively. The two-sided test checks the null hypothesis **H0:** $P_{\texttt{Rand}} = P_{\texttt{MaxWeight}}$ (with alternative $P_{\texttt{Rand}} \neq P_{\texttt{MaxWeight}}$), while the one-sided test checks null hypothesis **H0:** $P_{\texttt{Rand}} = P_{\texttt{MaxWeight}}$ (with alternative $P_{\texttt{Rand}} < P_{\texttt{MaxWeight}}$). We can reject *both* of these null hypotheses with $p \ll 0.01$. In light of the results presented in Table 1, these statistical test suggests `MaxWeight` achieves a small (∼ 5%) but significant improvement over `Rand` in terms of overall %MA. In the next set of statistical tests we compare each *day* of the experiment as a separate trial.

Daily Paired Comparison. Next we treat day of the experiment as a set of *paired measurements* of both $P_{\texttt{Rand}}$ and $P_{\texttt{MaxWeight}}$. For each day of the experiment (26 days in total) we calculate sample

[10]In Appendix E, we include a more detailed version of this table (Table 3) with results for 12 individual cities.

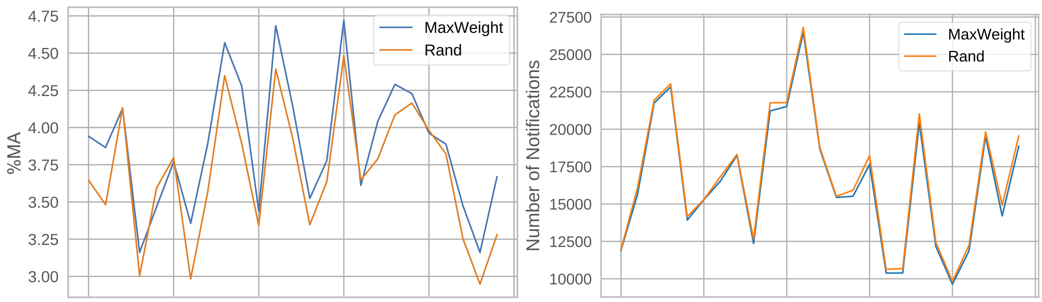


Fig. 5. (Left) Estimated probability of MA for the test group ($P_{\texttt{MaxWeight}}/100$) and and control group ($P_{\texttt{Rand}}/100$) for each day of the experiment, and (Right) total number of notifications sent to each group over 26 days.

estimates of $P_{\texttt{Rand}}$ and $P_{\texttt{MaxWeight}}$ – i.e., the 100 times the ratio of MAs to overall notifications. Note that donors are notified once every 14 days, meaning that the set of donors notified on any particular day is nearly disjoint from the donors notified on any other day of the experiment; for this reason we assume the measurements of $P_{\texttt{Rand}}$ and $P_{\texttt{MaxWeight}}$ on different days are independent. Figure 5 shows both the total number of notifications sent on each day of the experiment, with the daily measurement of $P_{\texttt{Rand}}$ and $P_{\texttt{MaxWeight}}$. Visually, `MaxWeight` nearly always achieves a higher %MA than `Rand` on any given day. To investigate whether this difference is significant, we treat these daily daily estimates as paired measurements of two separate distributions (i.e., the distribution of daily measurements of $P_{\texttt{MaxWeight}}$ and $P_{\texttt{Rand}}$). We use a two-sided Wilcoxon signed-rank test to check the null hypothesis **H0:** the median difference between $P_{\texttt{MaxWeight}}$ and $P_{\texttt{Rand}}$ is zero. We reject this null hypothesis ($p \ll 0.01$), further confirming that notification policy `MaxWeight` yields a higher MA rate than `Rand`.

8 DISCUSSION AND CONCLUSIONS

We introduce the problem of connecting blood donors with demand centers in a time-dependent setting, with uncertain demand. We formalize an objective for this problem, including the objectives of *efficiency* (maximizing the number of donations) and *fairness* for recipients. We propose a class of stochastic policies, to which we compare a realistic randomized baseline.

We test these policies in both a simulated donation environment (using real data), and a real donation system via the Facebook Blood Donation Tool. In simulations we see a clear tradeoff between the overall number of donations and fairness (see Figure 4); the particular tradeoff between these objectives depends on the notification policy used. Policy `MaxWeight` (which maximizes edge weight/expected donations) results in a 10-15% increase in the overall number of expected donations, compared to a random baseline (`Rand`). However `MaxWeight` also prioritizes certain recipients over others. In many cities, these simulations show that more than 50% of recipients are not notified by policy `MaxWeight` – presumably because these recipients are associated with lower edge weights. On the other hand, `Rand` nearly always sends a “fair” amount of notifications to each recipient, regardless of edge weight (see § 4.2 for a discussion of fairness). To mediate between the extremes of `Rand` and `MaxWeight`, we propose a class of stochastic policies (`LPMatch`(γ) and `LPMyopic`(γ)) In simulations these policies effectively control the balance between the overall expected number of donations and fairness for recipients, using parameter γ.

As a proof-of-concept we run an online experiment via the Facebook Blood Donation Tool, comparing notification policies `Rand` and `MaxWeight`. We find that `MaxWeight` results in about 5% more meaningful actions (a proxy for donations) than `Rand`. In relative terms this improvement seems small, however the implications are quite meaningful. This experiment investigated *one small improvement* to the notification strategy used by the Facebook Blood Donation Tool, i.e., whether the donor is notified about a nearby donation opportunity at random (`Rand`), or notified about a particular opportunity selected by a predictive model (`MaxWeight`). Several other modifications to the notification policy might yield similar improvements: for example by changing *how often* each donor is notified, by more carefully planning for *future donation needs*, or by tailoring notifications to each donor's unique preferences and values.

The potential impact of this work is considerable. Indeed, if our observed results generalize to the entire community of Facebook blood donors, then a 5% increase in donor action corresponds to at about $70,000$[11] *more* donors taking meaningful action toward donation. Even if few of these meaningful actions lead to actual donation, the increase is still substantial.

Before implementing these policies at a large scale in practice, it is important to understand their potential impacts on both blood donors and recipients. In this study impact on donors is minimal; the only difference between notification policies is *which donation opportunity* they are notified about. However our simulation results indicate that blood recipients may face significantly impacts from even a small change in notification policy. In simulations we find that notification policy `MaxWeight` ignores most recipients in many cities, presumably because these recipients are associated with low edge weights. In other words, `MaxWeight` tends to ignore recipients who are associated with low likelihood of meaningful action – which may include recipients in rural areas, or those with a limited Facebook presence. This observation is particularly troubling if low-weight recipients are *already* unlikely to recruit donors, which we expect is the case. Of course, this potential injustice is exactly the motivation for our stochastic policies `LPMatch`(γ) and `LPMyopic`(γ).

Blood donation is a global challenge, and has been the focus of many dedicated organizations and researchers for decades. In this paper we investigate a new opportunity to recruit and coordinate a massive network of blood donors and recipients, enabled by the widespread use of social networks. We formalize a matching problem around matching blood donors with recipients, and test these policies in both offline simulations and an online experiment using the Facebook Blood Donation Tool. Our findings suggest that a matching paradigm can significantly increase the overall number of donations, though it remains a challenge to do so while treating recipients equitably.

REFERENCES

[1] Rabeeh Ayaz Abbasi, Onaiza Maqbool, Mubashar Mushtaq, Naif R Aljohani, Ali Daud, Jalal S Alowibdi, and Basit Shahzad. 2018. Saving lives using social media: Analysis of the role of twitter for personal blood donation requests and dissemination. *Telematics and Informatics* 35, 4 (2018), 892–912.

[2] Turki Alanzi and Batool Alsaeed. 2019. Use of Social Media in the Blood Donation Process in Saudi Arabia. *Journal of Blood Medicine* 10 (2019), 417.

[3] Ross Anderson, Itai Ashlagi, David Gamarnik, and Yash Kanoria. 2017. Efficient dynamic barter exchange. *Operations Research* 65, 6 (2017), 1446–1459.

[4] Itai Ashlagi, Patrick Jaillet, and Vahideh H. Manshadi. 2013. Kidney Exchange in Dynamic Sparse Heterogenous Pools. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*. 25–26.

[5] Vidar Bosnes, Magne Aldrin, and Hans Erik Heier. 2005. Predicting blood donor arrival. *Transfusion* 45, 2 (2005), 162–170.

[6] Anna Bárbara Carneiro-Proietti, Ester C Sabino, Divaldo Sampaio, Fernando A Proietti, Thelma T Gonçalez, Cláudia DL Oliveira, João E Ferreira, Jing Liu, Brian Custer, George B Schreiber, et al. 2010. Demographic profile of blood donors at three major Brazilian blood centers: results from the International REDS-II study, 2007 to 2008. *Transfusion* 50, 4 (2010), 918–925.

[7] Kathleen Chell, Tanya E Davison, Barbara Masser, and Kyle Jensen. 2018. A systematic review of incentives in blood donation. *Transfusion* 58, 1 (2018), 242–254.

[11] About 0.2% of all blood donors registered with the Facebook Blood Donation Tool ((https://about.fb.com/news/2019/06/us-blood-donations/), corresponding to the absolute increase in meaningful action rate reported in Table 1.

[8] Mark Cieliebak, Stephan Eidenbenz, Aris Pagourtzis, and Konrad Schlude. 2008. On the Complexity of Variations of Equal Sum Subsets. *Nord. J. Comput.* 14, 3 (2008), 151–172.

[9] Paul D Cumming, Kenneth E Kendall, C Carl Pegels, John P Seagle, and John F Shubsda. 1976. A collections planning model for regional blood suppliers: description and validation. *Management Science* 22, 9 (1976), 962–971.

[10] John P Dickerson, Karthik A Sankararaman, Aravind Srinivasan, and Pan Xu. 2018. Allocation problems in ride-sharing platforms: Online matching with offline reusable resources. In *AAAI Conference on Artificial Intelligence (AAAI)*.

[11] Mary Dillon, Fabricio Oliveira, and Babak Abbasi. 2017. A two-stage stochastic programming model for inventory management in the blood supply chain. *International Journal of Production Economics* 187 (2017), 27–41.

[12] Lester E Dubins and Edwin H Spanier. 1961. How to cut a cake fairly. *The American Mathematical Monthly* 68, 1P1 (1961), 1–17.

[13] Jon Feldman, Aranyak Mehta, Vahab Mirrokni, and S. Muthukrishnan. 2009. Online Stochastic Matching: Beating 1-1/e. In *Symposium on the Foundations of Computer Science (FOCS)*.

[14] Simone A Glynn, Michael P Busch, George B Schreiber, Edward L Murphy, David J Wright, Yongling Tu, Steven H Kleinman, NHLBI REDS Study Group, et al. 2003. Effect of a national disaster on blood supply and safety: the September 11 experience. *Jama* 289, 17 (2003), 2246–2253.

[15] Simone A Glynn, Alan E Williams, Catharie C Nass, James Bethel, Debra Kessler, Edward P Scott, Joy Fridey, Steven H Kleinman, George B Schreiber, and Retrovirus Epidemiology Donor Study. 2003. Attitudes toward blood donation incentives in the United States: implications for donor recruitment. *Transfusion* 43, 1 (2003), 7–16.

[16] Gaston Godin, Mark Conner, Paschal Sheeran, Ariane Bélanger-Gravel, and Marc Germain. 2007. Determinants of repeated blood donation among new and experienced blood donors. *Transfusion* 47, 9 (2007), 1607–1615.

[17] Gaston Godin, Paschal Sheeran, Mark Conner, Marc Germain, Danielle Blondeau, Camille Gagné, Dominique Beaulieu, and Herminé Naccache. 2005. Factors explaining the intention to give blood among the general population. *Vox sanguinis* 89, 3 (2005), 140–149.

[18] Gagan Goel and Aranyak Mehta. 2008. Online budgeted matching in random input models with applications to Adwords. In *Proceedings of ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 982–991.

[19] Gagan Goel and Pushkar Tripathi. 2012. Matching with our eyes closed. In *Symposium on the Foundations of Computer Science (FOCS)*. IEEE, 718–727.

[20] Yue Guan. 2018. When Voluntary Donations Meet the State Monopoly: Understanding Blood Shortages in China. *The China Quarterly* 236 (2018), 1111–1130.

[21] Chien-Ju Ho and Jennifer Wortman Vaughan. 2012. Online task assignment in crowdsourcing markets. In *AAAI Conference on Artificial Intelligence (AAAI)*.

[22] Patrick Jaillet and Xin Lu. 2014. Online stochastic matching: New algorithms with better bounds. *Mathematics of Operations Research* 39, 3 (2014), 624–646.

[23] R. M. Karp, U. V. Vazirani, and V. V. Vazirani. 1990. An Optimal Algorithm for On-line Bipartite Matching. In *Proceedings of the Annual Symposium on Theory of Computing (STOC)*. 352–358.

[24] Korina Katsaliaki and Sally C Brailsford. 2007. Using simulation to improve the blood supply chain. *Journal of the Operational Research Society* 58, 2 (2007), 219–227.

[25] Meghna Lowalekar, Pradeep Varakantham, and Patrick Jaillet. 2018. Online spatio-temporal matching in stochastic and dynamic domains. *Artificial Intelligence* 261 (2018), 71–112.

[26] Aranyak Mehta. 2012. Online matching and ad allocation. *Theoretical Computer Science* 8, 4 (2012), 265–368.

[27] Steven A Melnyk, Mark Pagell, Gail Jorae, and A Spencer Sharpe. 1995. Applying survival analysis to operations management: Analyzing the differences in donor classes in the blood donation process. *Journal of Operations Management* 13, 4 (1995), 339–356.

[28] Andres F Osorio, Sally C Brailsford, and Honora K Smith. 2015. A structured review of quantitative models in the blood supply chain: a taxonomic framework for decision-making. *International Journal of Production Research* 53, 24 (2015), 7191–7212.

[29] Andres F Osorio, Sally C Brailsford, Honora K Smith, Sonia P Forero-Matiz, and Bernardo A Camacho-Rodríguez. 2017. Simulation-optimization model for production planning in the blood supply chain. *Health Care Management Science* 20, 4 (2017), 548–564.

[30] Sofia Ouhbi, José Luis Fernández-Alemán, Ambrosio Toval, Ali Idri, and José Rivera Pozo. 2015. Free blood donation mobile applications. *Journal of medical systems* 39, 5 (2015), 52.

[31] Pascale Reich, Paula Roberts, Nancy Laabs, Artina Chinn, Patrick McEvoy, Nora Hirschler, and Edward L Murphy. 2006. A randomized trial of blood donor recruitment strategies. *Transfusion* 46, 7 (2006), 1090–1096.

[32] George B Schreiber, UK Sharma, DJ Wright, SA Glynn, HE Ownby, Y Tu, G Garratty, J Piliavin, T Zuck, R Gilcher, et al. 2005. First year donation patterns predict long-term commitment for first-time donors. *Vox sanguinis* 88, 2 (2005), 114–121.

[33] B Nilsson Sojka and Peter Sojka. 2008. The blood donation experience: self-reported motives and obstacles for donating blood. *Vox sanguinis* 94, 1 (2008), 56–63.

[34] Ariane Sümnig, Martin Feig, Andreas Greinacher, and Thomas Thiele. 2018. The role of social media for blood donor motivation and recruitment. *Transfusion* 58, 10 (2018), 2257–2259.

[35] Anne Van Dongen, Robert Ruiter, Charles Abraham, and Ingrid Veldhuizen. 2014. Predicting blood donation maintenance: the importance of planning future donations. *Transfusion* 54, 3pt2 (2014), 821–827.

[36] Rakesh V Vohra. 2011. *Mechanism design: a linear programming approach*. Vol. 47. Cambridge University Press.

[37] World Health Organization WHO. 2017. Blood safety and availability. https://www.who.int/news-room/fact-sheets/detail/blood-safety-and-availability [Online; accessed 22-July-2004].

[38] PLH Yu, KH Chung, CK Lin, JSK Chan, and CK Lee. 2007. Predicting potential drop-out and future commitment for first-time donors based on first 1· 5-year donation patterns: the case in Hong Kong Chinese donors. *Vox sanguinis* 93, 1 (2007), 57–63.

[39] Shan Yuan, Shelley Chang, Kasie Uyeno, Gay Almquist, and Shirong Wang. 2016. Blood donation mobile applications: are donors ready? *Transfusion* 56, 3 (2016), 614–621.

[40] Behzad Zahiri and Mir Saman Pishvaee. 2017. Blood supply chain network design considering blood group compatibility under uncertainty. *International Journal of Production Research* 55, 7 (2017), 2013–2033.

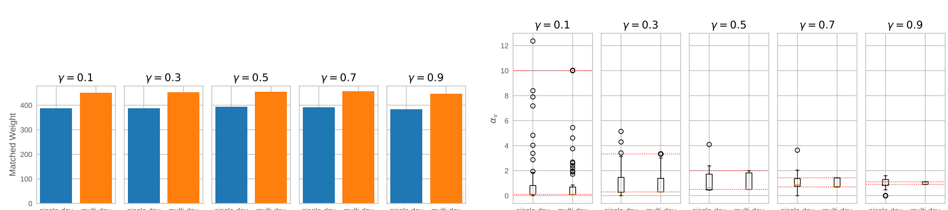
A COMPARING THE MYOPIC AND OMNISCIENT LPS

In this section we compare the expected behavior of the LP-relaxation used by policy `LPMyopic`(). In particular, we investigate whether using a longer or shorter planning horizon H has any impact on the objective value or constraints. We refer to the LP relaxation with an H-day horizon (i.e., Problem 3) as `H-day`.

To compare these LPs we construct a 5-day matching scenario using donor and recipient data from a single city, with 29,646 donors, 80 recipients, and 36,700 edges. Donors can be notified once every $K = 14$ days, and we consider a rolling time horizon of $H = 5$ steps (i.e., the complete time horizon). We randomly assign each donor a prior-notification date, uniformly distributed between $-K$ and 0 days; this reflects the fact that donors arrive on a "rolling" schedule, not all at once. However `1-DAY` can only use knowledge on each given day. To simulate the `1-DAY` approximation, we solve the single-day LP (Problem **??**) for each day sequentially; on each day, all available donors are matched. Note that both `N-DAY` and `1-DAY` match exactly the same donors over the complete time horizon, but at possibly different times. There are two quantities of interest in comparing these policies: (a) the objective value (i.e., the expected number of meaningful actions), and (b) violation of the fairness constraints. Recall that our notion of proportional fairness pertains to each recipient's utility over the *entire time horizon* $\mathcal{T}$. These constraints are explicit in the `N-DAY`, while `1-DAY` includes constraints *only for each day*. To investigate these constraints we calculate the proportional outcome for each recipient α_v, i.e., the expected scaled utility for recipient v divided by the average scaled utility over all recipients. In these experiments α_v is calculated the same way as in previous sections, only using *fractional* decision variables $x^*_{et} \in [0, 1]$. Recall that our proportional fairness constraints require that, for each recipient v,

$$\gamma \leq \alpha_v \leq 1/\gamma.$$

To visualize the violation of these fairness constraints, we plot α_v for both `1-DAY` and `N-DAY`.



(a) (b)

Fig. 6. (a) Objective value (total matched weight) of single-day LP (`1-DAY`) and multi-day LP (`N-DAY`), (b) proportional outcomes α_v for each recipient v; red dotted lines indicate fairness constraints.

Figure 6a shows the objective values for both LP solutions, for $\gamma \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$. The `1-DAY` approximation consistently achieves a lower objective value than `N-DAY` (by about 15%). This is expected: `N-DAY` can match a donor on *any day* after they arrive; in contrast, `1-DAY` must match each donor immediately. Surprisingly, the total objective both both formulations is nearly the same constant for each γ.

Figure 6b shows boxplots of the ratio $U_v/\overline{U}$ for each recipient; red dotted lines indicate the upper and lower fairness constraints (i.e., γ and $1/\gamma$). As expected, `N-DAY` never violates these constraints (i.e., the boxplots are completely within the red dotted lines), while `1-DAY` violates these constraints.

Table 2 includes the total objective value and number of constraint violations for both LP solutions.

Table 2. Comparison of multi-day LP solution (N-DAY) to single-day approximation (1-DAY) for various γ. Objective value (Total Weight) and number of constraint violations (# Violations) are shown.

	$\gamma = 0.1$	$\gamma = 0.3$	$\gamma = 0.5$	$\gamma = 0.7$	$\gamma = 0.9$
N-DAY					
Total Weight	450.0	452.6	454.1	455.5	445.1
# Violations	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
1-DAY					
Total Weight	386.0	385.9	392.2	390.6	383.7
# Violations	3 (4%)	7 (11%)	17 (25%)	27 (39%)	33 (71%)

The 1-DAY approximation consistently achieves a lower objective value than N-DAY (by about 15%), though the total objective is relatively constant for different values of γ.

Of course for this reason, 1-DAY is a more practical policy, as it makes no assumptions about the future.

We might expect that policy 1-DAY

the difference between the fractional matching solution produced by the single-day and multi-day LP formula. Specifically, we compare (a) the objective value (total weight, i.e., expected number of meaningful actions), and (b) violation of the fairness constraints. Recall that our notion of proportional fairness pertains to each recipient's utility over the entire time horizon $\mathcal{T}$. These constraints are explicit in the multi-day LP formulation (Problem 3); however the single-day formulation includes fairness constraints for only one day.

We compare

two proportionally fair policies with different time horizons. We aim to demonstrate that a single-day policy (i.e., γ-FAIR) both (a) achieves about the same matched weight as a K-day policy, and (b) does not significantly violate fairness constraints.

B PROOF OF COMPETITIVE RATIO

In this section we bound the worst-case performance our randomized policy γ-FAIR, compared with an *offline optimal* notification policy. With some abuse of notation, let $E[\gamma\text{-FAIR}]$ denote the expected matching weight of policy γ-FAIR, given the distribution of recipient arrivals p_{vt} and

$$
\begin{array}{lll}
\max & \sum_{t \in T^H} \sum_{e \in E(t)} w_{et} x_{et} & \\
\text{s.t.} & x_{et} \in [0,1] & \forall e \in E(t)\ t \in T^H \\
& U^s_{vt} \in \mathbb{R} & \forall v \in V(t), t \in T^H \\
& x_{et} \leq p_{vt} & \forall v \in V(t), e \in E_{:v}(t), t \in T^H \\
& \sum_{t'=t}^{t+k} \sum_{e \in E_{u:}(t)} x_{et} \leq 1 & \forall u \in U, t = \{1, \dots, T^H - k\} \\
& U^s_v = \frac{1}{m_v} \sum_{t \in T^H} \sum_{e \in E_{:v}(t)} x_{et} w_{et} & \forall v \in V_H \\
& U^s_v \leq \frac{1/\gamma}{|V_H|} \sum_{v' \in V_H} U^s_{v'} & \forall v \in V_H \\
& U^s_v \geq \frac{\gamma}{|V_H|} \sum_{v' \in V_H} U^s_{v'} & \forall v \in V_H
\end{array}
\tag{3}
$$

C MIN COST FLOW FOR A LIMITED PLANNING HORIZON

In this section we assume that $T \leq K$, which means that every donor can be matched at most once. This assumption can be leveraged to make certain variants of our problem substantially more tractable. Furthermore, we assume that we are given upper and lower bounds u_v, l_v on demand for each recipient v. With these assumptions, an optimal assignment for the model can be computed in polynomial time (and efficiently in practice) via *min-cost flow*; this is easily seen by noting that the problem for a single time step is an instance of the *transportation problem*. Multiple timesteps can be handled by replicating each recipient T times, where replica t denotes the recipient receiving donors on the t'th day (this works because $T \leq K$ and thus each donor can be used at most once).

Formally, we construct the network flow problem as follows: each donor $u \in U$ is instantiated with a source node s_u with outgoing flow of 1. Each recipient $v \in V$ is represented by nodes $n_{v1}, \ldots, n_{vT}$, where n_{vt} represents the donors assigned to v at day t (we can of course leave out pairs v, t such that v is unavailable on day t, e.g. for events of limited duration). Each n_{vt} is connected to the single sink node, and the connecting edge has upper and lower bounds u_v, l_v on capacity, and cost 0, we let E_{vs} denote the edges from recipient v to the sink. For each edge in the original graph $e \in E$ we add edges $e_1, \ldots, e_T$, with e_t representing using the edge e on day t; each e_t has cost $-w_e$ and infinite capacity.

A solution to the min-cost flow model can be computed with the following LP

$$
\begin{array}{lll}
\min & \sum_{t \in \mathcal{T}} \sum_{e \in E} -w_e x_{et} & \\
\text{s.t.} & \sum_{t \in \mathcal{T}: t-K \leq t' \leq t} \sum_{e \in E_{u:}} x_{et'} \leq 1 & \forall u \in U, t \in \mathcal{T} \\
& l_v \leq x_{et} \leq u_v & \forall v \in V, e \in E_{vs}, t \in \mathcal{T} \\
& 0 \leq x_{et} \leq 1 & \forall e \in E, t \in \mathcal{T}
\end{array}
\tag{4}
$$

Since each capacity is integral (or infinite), we know that an optimal integer solution exists, due to the well-known property that basis solutions of the min-cost flow LP are integral in this case (see e.g. [36]). Since $T \leq K$ we know that each donor can be assigned at most one time, and thus the above construction is valid.

D PROOF OF NP-HARDNESS

Proof. Proof of Theorem 4.2. The proof is by reduction from the k-EQUAL-SUM-SUBSET problem: given a multiset $\mathcal{S}$ of $x_1, \ldots, x_n$ positive integers, are there k nonempty disjoint subsets $S_1, \ldots, S_k \subset \mathcal{S}$ such that $sum(S_1) = \ldots = sum(S_k)$. This problem is NP-hard for any fixed $k > 1$, but strongly NP-hard when k varies as a function of n and $k = \Omega(n)$ [8].

Given an instance of k-EQUAL-SUM-SUBSET we construct an instance of fair donation matching as follows: we add k recipients, one for each subset, and n donors, one for each integer x_i. Each donor i has edge weight x_i to every recipient. First consider the case where $\gamma = 1$. In that case since the proportional allocation awards the same utility to every recipient, we must find an allocation that gives exactly the same utility to every recipient. A solution that allocates every donor now corresponds exactly to an equal-sum partitioning. If such a solution exists, then it is the solution that maximizes overall donations, since all weights are positive.

Finally, consider the case where $\gamma < 1$. A solution to our reduction above, if there exists a solution to k-EQUAL-SUM-SUBSET, would give every recipient utility $\frac{1}{3} \sum_i x_i$. Now we add an auxiliary recipient that has a donor with an edge only to them, with weight $\frac{1}{3\gamma} \sum_i x_i$. Now a γ-proportionally-fair allocation that allocates any utility to the auxiliary recipient requires all recipients to get utility $\frac{1}{3} \sum_i x_i$, which is possible iff there is a solution to k-EQUAL-SUM-SUBSET. □

Table 3. Online Experiments - Number of notifications (#Notifs) and meaningful actions (#MA), over the online experiment. Notifications are separated into those sent to donors with only one compatible recipient (1R), and those sent to donors with two or more compatible recipients (≥2R). Wilson score intervals are for the percentage of notifications that lead to MA are presented as $C \pm R/2$, where the 95% confidence interval is $[C - R/2, C + R/2]$. The top row shows results for the entire experiment, while the following 12 rows show results from each of the cities corresponding to Figure 4 (a small sampling of the overall experiment).

City Number	Notif. Group	Control (Rand)			Test (MaxWeight)		
		#MA	#Notifs	%MA	#MA	#Notifs	%MA
Experiment	1R	10,534	215,544	4.7 ± 0.1	10,755	214,841	4.8 ± 0.1
	≥2R	17,479	469,887	3.6 ± 0.1	18,030	459,708	3.8 ± 0.1
1	1R	7	97	8.3 ± 5.0	6	95	7.6 ± 4.8
	≥2R	14	255	5.8 ± 2.7	16	256	6.5 ± 2.8
2	1R	2	62	5.8 ± 4.9	1	61	4.4 ± 4.2
	≥2R	37	1,822	2.1 ± 0.6	38	1,817	2.1 ± 0.7
3	1R	5	63	9.6 ± 6.5	6	50	13.2 ± 8.2
	≥2R	36	495	7.1 ± 2.2	35	496	6.9 ± 2.1
4	1R	14	200	7.3 ± 3.4	10	192	5.8 ± 3.1
	≥2R	54	1,109	4.8 ± 1.2	62	1,139	5.3 ± 1.3
5	1R	4	10	33.2 ± 21.5	1	12	17.3 ± 16.0
	≥2R	17	256	6.8 ± 2.9	13	290	4.9 ± 2.3
6	1R	3	55	8 ± 6.2	3	47	9.1 ± 7.1
	≥2R	16	247	6.7 ± 2.9	22	224	9.6 ± 3.6
7	1R	1	57	4.7 ± 4.4	0	45	3.9 ± 3.9
	≥2R	71	2,099	3.4 ± 0.8	71	2,070	3.4 ± 0.8
8	1R	3	25	15.5 ± 11.7	1	23	10.5 ± 9.8
	≥2R	14	342	4.4 ± 2.1	31	381	7.9 ± 2.6
9	1R	11	148	7.9 ± 4.0	7	140	5.9 ± 3.6
	≥2R	141	3,755	3.7 ± 0.6	170	3,949	4.2 ± 0.6
10	1R	0	33	5.2 ± 5.2	3	45	9.5 ± 7.3
	≥2R	20	262	7.7 ± 3.0	20	289	7.0 ± 2.8
11	1R	313	5,467	5.4 ± 0.6	318	5,376	5.6 ± 0.6
	≥2R	14	268	5.6 ± 2.6	17	218	7.9 ± 3.4
12	1R	48	668	6.9 ± 1.8	39	646	5.9 ± 1.7
	≥2R	70	854	7.8 ± 1.7	66	843	7.4 ± 1.7

E ADDITIONAL REAL-WORLD ONLINE EXPERIMENTS

Table 3 below is an extended version of Table 1 from Section 7.1 in the main body of the paper. Both tables summarize results from our real-world, online experiment; additionally, Table 3 gives results for 12 individual cities.