Workforce Configuration in Charity Settings: A Forward-looking Approach

(Authors' names blinded for peer review)

Problem definition: Serving as the primary workforce for many charitable organizations, volunteers represent a unique and complex labor pool. Not only are they often unreliable, but they also exhibit considerable heterogeneity in performance and affinity to the organization. Further, many volunteers engage not merely to contribute but also to partake in a *volunteering experience*, which has the potential to transform them into future donors. Acknowledging these distinct features, this paper aims to present a forward-looking optimization model to refine volunteer scheduling. **Methodology/results:** Building on our field study and insights from existing literature, we propose a model that accounts for the diversity among volunteers, mitigates both understaffing and overstaffing costs, and explicitly correlates individual contributions in both time and monetary donations. We provide analytical solutions when the charity can reliably estimate distributions from data. In cases where data is limited or uncertain, we suggest a distribution-free method to offer actionable insights for managerial decision-making. **Managerial implications:** At a high level, by viewing volunteers as potential donors, the optimal staffing strategy balances the charity's need to fulfill its labor requirements against the workers' utility of the volunteering experience, since the latter influences whether the volunteer is transformed into a future donor. Our findings reveal that the charity in our study could avert substantial losses by adopting this integrative approach, thereby challenging the conventional organizational structures in charities that compartmentalize volunteer and donor management. We also distill several managerial insights, such as identifying scenarios where: (i) the *episodic* volunteers are the preferred group to be invited, and (ii) enhancements in labor performance could paradoxically diminish the charity's overall utility. Based on our field study and computational experiments, we have identified advantages and disadvantages associated with various policies. However, we also demonstrate that building a robust data infrastructure can markedly improve volunteer management performance. We conclude by offering an Excel-based decision support tool and a decision-tree framework designed to navigate optimal policies within the constraints of operational realities.

Keywords: workforce management; volunteer labor scheduling; optimization.

---

1. Background

In 2020, despite the formidable challenges posed by the Covid-19 pandemic, an estimated 23% of the U.S. population–equivalent to more than 60 million individuals–volunteered their time and skills in charitable organizations. These volunteers collectively logged an impressive 4.1 billion hours, representing an economic value of a staggering $122.9 billion (AmeriCorps 2023). Similarly active volunteerism is observed outside of the U.S. For instance, out of the 5.2 million Australians who volunteered their time in 2006, a remarkable 84% collectively contributed 623 million hours to the Australian non-profit sector, corresponding to an estimable 15 billion Australian dollars (Productivity Commission 2008). Volunteers play a pivotal role in a plethora of administrative and operational tasks, constituting a cornerstone of the labor force for many charitable institutions. Emerging research in volunteer management has illuminated various dimensions of this field, from the nuanced relationship between volunteers' time and monetary donations (Brown et al. 2019), to innovative strategies for enhancing volunteer motivation (Gage and Thapa 2012; Beder and Fast 2008; Hustinx et al. 2008), to volunteer retention (Cnaan and Handy 2005; Hyde et al. 2016), to mechanisms for improving volunteer retention (Dwiggins-Beeler et al. 2011), and even to the development of optimal policies for matching a volunteer's preferences with suitable opportunities (Brudney and Meijs 2009; Manshadi and Rodilitz 2022). However, what remains conspicuously absent is a dedicated investigation into the volunteer staffing question, particularly one that takes into account the idiosyncratic characteristics intrinsic to managing a volunteer workforce. The goal of this study is to fill this research gap.

To address this issue, one must first understand the unique characteristics that define volunteer management in charitable organizations, characteristics that are markedly different from the features of workforce management in commercial enterprises. Contrary to the conventional wisdom that volunteers are ubiquitously available, uniformly motivated, and universally capable, the reality is far more nuanced. Volunteers exhibit a considerable heterogeneity, not only in the strength of their affiliation with the charity but also in their reliability and performance levels. For instance, volunteers with tenuous ties to the charity are often less dependable in their turnout and may be less effective in task execution owing to their lack of experience. Such volunteers can inadvertently become a costly labor supply for the charity. Take, for example, a volunteer who fails to appear for a "meal packaging" event. The absence engenders dual costs for the charity: a shortage cost due to unmet demand and an obsolescence cost attributable to the waste of unused raw materials.

In practice, volunteers generally fall into two categories: *formal* and *episodic* (Hustinx et al. 2008). Formal (also known as *regular*) volunteers are those who display a long-term commitment and a profound dedication to a specific cause. They serve with regularity, often resembling paid staff in terms of reliability and scheduled attendance. In contrast, episodic volunteers make up the majority of the volunteer workforce (Low et al. 2007; Cnaan et al. 2022). Their involvement is sporadic and typically confined to specific periods. These individuals are drawn to short-term flexible tasks and are less motivated by altruistic and social factors (Hustinx and Lammertyn 2003; Beder and Fast 2008; Cnaan et al. 2022). For instance, our field study at a local charity (the Society of St. Vincent de Paul in Phoenix) revealed a striking challenge: the no-show rate for episodic volunteers averages around 30%. Most absentees either fail to inform coordinators of their absence or do so too late to secure a replacement. Episodic volunteers' high rate of unreliability leads to significant understaffing, resulting in unmet demands and increased operational costs for the charity (Hyde et al. 2016; Ata et al. 2019). While the ideal volunteer base might consist of a dedicated cadre of formal volunteers, this group is not only limited in number but also appears to be in decline nationally (Brudney and Meijs 2009); In today's fast-paced world, many people are too busy to make long-term commitments to a single organization or cause, preferring instead brief engagements (Cnaan et al. 2022). Consequently, managers often find themselves needing to augment their core team with episodic volunteers to adequately meet operational demands. However, overcompensating for potential absenteeism by inviting an excess of volunteers is not a viable solution (Ellis 2007). Such overstaffing risks creating an environment where volunteers feel superfluous, which in turn could lessen their eagerness to contribute in subsequent instances (Smith 1998; Sampson 2006; Dwiggins-Beeler et al. 2011). Vecina et al. (2012) show that volunteer satisfaction is projected to influence individuals' intention to continue volunteering, at least in the short term. Our analysis of volunteer surveys from the charity corroborates this concern; numerous volunteers expressed dissatisfaction, stating they "*had little work to do and left early due to sufficient staffing.*" This sentiment aligns with a survey analysis by Sampson (2006), which indicates that both the underutilization and overextension of volunteer labor can discourage future participation.

Another significant challenge arises when episodic volunteers, spontaneously invite friends or family to join them, often without prior notification to the charity. This unscheduled influx can lead to an overstaffing issue, making the event appear disorganized. However, turning away these *uninvited* volunteers is generally not an option, discourage individuals from making future contributions in any capacity (Daniels and Valdés 2021). Moreover, as one charity manager put it, "*Volunteering work is an experience we offer to people, and our aim is to extend this experience to as many individuals as possible. Yet, we don't want them to arrive only to discover there is little to be done.*" As a result, episodic volunteers introduce an element of volatility into the staffing equation; they may either fail to appear as scheduled or unexpectedly arrive with additional, unsolicited volunteers in tow.

Second, contrary to conventional wisdom, the contributions of volunteers extend beyond mere labor. Studies show that the relationship between individuals' time and monetary donations is complementary (Brown and Lankford 1992; Cappellari et al. 2011), and even suggest that volunteering can actually boost financial contributions (Apinunmahakul et al. 2009). In essence, volunteering not only raises awareness about the charity's needs but also enhances transparency, thereby building trust between the individual and the organization (Parsa et al. 2022). Such events serve as crucial touchpoints that solidify the bond between volunteers and the charity, having a significant impact on future donations (Olsen and Eidem 2003; Feldman 2010; Bekkers and Wiepking 2011). Many volunteers prefer to "test the waters" by volunteering before they commit to financial donations (Fritz 2019; Dietz and Keller 2016). This perspective is corroborated by experiments showing that individuals who first consider donating their time are more likely to also donate money (Liu and Aaker 2008). Recent laboratory experiments have demonstrated that due to the effect of moral consistency, participants assigned to volunteering tasks are more likely to donate, and in larger amounts, compared to those who did not serve as volunteers (Authors 2023). These findings align with field studies; for example, a survey conducted by Fidelity Charitable (2014) shows that over half of volunteers indicate that their volunteering experience leads them to make financial contributions. Another survey found that 87% of individuals who support charities donate both time and money, and 43% donate money to the same charities where they volunteer (Fidelity Charitable 2014). This illuminates the dual role of a volunteer: as a *producer* of social welfare and also as a *customer* of the volunteering experience. The satisfaction derived by the volunteer-customer during their engagement not only influences their immediate contributions but also shapes their future relationship with the charity, including potential financial donations (Miller et al. 1990; Clary et al. 1998; Dwiggins-Beeler et al. 2011). For the long-term sustainability of the charity, it is imperative to reconcile these dual roles when planning staffing needs. Yet, the common organizational structure in charities often silos volunteer management and donor management into distinct functions. Specifically, volunteer program managers focus on task design and volunteer recruitment, while development managers oversee fundraising initiatives and donor relationships. This compartmentalized approach misses the opportunity to leverage the synergistic potential between volunteerism and donorship.

Third, the composition of volunteers at each event plays a pivotal role in enhancing their engagement. Generally, the relationship between episodic volunteers and the organizations they serve is often transient and task-oriented, lacking the psychological contract commonly found with formal volunteers (Vantilborgh et al. 2011). Unlike formal volunteers, who often develop expectations and sustain dyadic relationships with the host organization, episodic volunteers tend to engage solely for the completion of a specific task, often disappearing thereafter with a likelihood of not returning (Cnaan et al. 2022). This nature of engagement leads to high turnover rates among episodic volunteers, who are typically drawn to enjoy the occasion or event (Hyde et al. 2016). Formal volunteers often represent a homogeneous group, sharing similar identities, beliefs, and motivations towards a specific social cause (Charness and Chen 2020). Research indicates that such homogeneous groups are more effective in contributing to public goods compared to their heterogeneous counterparts (Burlando and Guala 2005; Gachter and Thoni 2005; Ai et al. 2016). Moreover, homogeneity fosters stronger group cohesion (Bugen 1977; Lieberman et al. 2005), which in turn cultivates a sense of collective psychological ownership among members (Pierce and Jussila 2010). This heightened sense of ownership enhances the likelihood of future donations (Peck et al. 2021; Jami et al. 2021). Additionally, the frequency of interactions among formal volunteers fosters a robust group identity (Fraser et al. 2009; Gray and Stevenson 2020). As such, a formal volunteer's gratification in contributing to a charity's mission–and consequently, their propensity for future donations–is influenced not just by their individual commitment, but also by the endogenous decision-making processes related to team composition. For instance, survey data from our observed charity reveals a pronounced preference among formal volunteers for working within their usual groups.[1] As a result, a formal volunteer's pleasure in contributing to the charity's mission – and consequently, their inclination for future donations – is partly shaped by the internal dynamics of team composition. In light of this, we posit that formal volunteers are likely to make more substantial donations when grouped with their regular, like-minded counterparts. On the other hand, episodic volunteers appear to exhibit a more laissez-faire attitude toward team composition.

The aim of this study is to craft a streamlined volunteer management model that accounts for the unique complexities of volunteer scheduling, ultimately enhancing a charity's long-term utility. To ascertain the optimal number of volunteers and the ideal team composition for each event, we introduce a framework that accommodates volunteer heterogeneity, turnout uncertainty, both understaffing and overstaffing costs, and the dual role of volunteers as both labor and potential future donors. In addition, we take into account the composition of volunteer teams. Despite the additional complexity leading to a non-convex model structure, we were able to derive a closed-form expression for the optimal staffing plan. This allows us to offer actionable insights for improving charity processes. We show that reducing volunteer turnout uncertainty is universally beneficial for charities. This could be achieved, for instance, by sending tailored messages that underscore the significance of the task at hand, thereby enhancing the value-based aspects of volunteers' psychological contracts (Vantilborgh et al. 2012; De La Torre Pacheco et al. 2023). Intriguingly, while conventional wisdom might suggest that improving the efficiency of reliable volunteers through training programs would be advantageous, our model reveals this is not always the case. This counterintuitive finding arises because more efficient volunteers lessen the total number of volunteers needed for an event, which in turn could reduce monetary donations due to fewer volunteers experiencing the activity. Therefore, any efficiency-boosting training must be harmonized with a redesign of the volunteer event to ensure adequate tasks for a comparably sized volunteer pool. Our study challenges traditional charity structures by introducing a volunteer management model that explicitly recognizes the multifaceted role of volunteers.

Moreover, we supplement our model with a series of numerical experiments, conducted in partnership with a local charity in Phoenix, Arizona, as well as a nonprofit consulting firm.[2] One critical insight garnered from these collaborations is the prevalent issue of data quality, especially in the realm of volunteer management. This challenge partly stems from charities' relentless efforts to minimize overhead costs, leaving little room for infrastructural investments (Parsa et al. 2022). Further, in practice, many charities are also reticent to monitor volunteer attendance and performance rigorously. For instance, a manager at our focus charity noted, "*We refrain from* grading *our volunteers*

[1] Comments such as "*I prefer to work with the usual group*," "*I love working with my regular fellow volunteers*," and "*It feels like the volunteers are a second family*" underscore this sentiment.

[2] AmPhil: https://amphil.com/

because it is a sensitive subject. Our concern is that such evaluations will turn them away." This reluctance to gather volunteer data results in a lack of automated and consistent data collection methods. Our examination revealed that approximately 40% of this particular charity's data records contained some degree of error.[3] In light of these challenges, we extended our model by adopting a distributionally robust optimization approach, which we solve numerically. The method we propose for solving the robust model is presented in the Appendix. Our numerical experiments demonstrate that implementing the proposed closed form policy generates a median increase of 6.41% in the combined value of labor and donations. Alternatively, deploying the distributionally robust solution elevates the total utility by a median of 8.87%. A comparative analysis of various models reveals a paradigm shift in the role of volunteers when a charity incorporates future donations into its planning calculus. Traditional volunteer planning models, which disregard the potential for future donations, prioritize highly reliable volunteers who have strong affiliations with the charity for staffing events. Less reliable volunteers are only called upon in scenarios of labor shortages. In contrast, our model, which accounts for volunteers as prospective donors, suggests that less reliable volunteers might actually be preferable in certain contexts. This counterintuitive finding emerges when these less reliable volunteers place a high intrinsic value on the volunteering experience, translating into increased monetary donations that outweigh the expected shortfall in labor contributions due to their unreliability.

Labor scheduling remains a pivotal subject in operations management, primarily driven by the profound impact of labor costs on operational expenditures (Van den Bergh et al. 2013; Smilowitz et al. 2013); Industries like healthcare and service desks witness labor salaries being responsible for over half of their operational overheads (Villarreal et al. 2015). Thus, even marginal enhancements in labor productivity and morale can usher in notable savings, curbing staff attrition. Taking a magnifying lens to commercial settings, Van den Bergh et al. (2013) offer an exhaustive literature survey, illustrating a thorough review of various employment settings, including full-time and part-time positions, flexible schedules, and workers' preferences for teamwork or specific shifts. Berman

[3] This challenge becomes even more convoluted in larger food banks offering a diverse array of volunteering tasks, as attendance recording processes can differ markedly across various programs. For instance, the charity we studied provides approximately 44 distinct volunteering tasks and adapts this list based on evolving needs. While some programs employ dedicated staff to meticulously track volunteer attendance, others leave the responsibility of check-in and check-out to the volunteers themselves. Feedback from volunteer surveys revealed a significant pain point: volunteers find the process of clocking in and out to be both redundant and cumbersome. Many expressed that this administrative chore makes them feel as though they are *working* rather than *volunteering*, undermining the altruistic essence of their involvement. Consequently, we observed numerous manual data entry errors in the charity's attendance records. It is not uncommon for the attendance record of a volunteer involved in a recurring event for several months to indicate either a "no-show" or to be left blank. At times, volunteer managers discover these discrepancies only at the conclusion of a program, at which point they might attempt to rectify the record by inputting an "estimated total volunteering hours" or by appending comments to account for the missing entries. Unfortunately, such post-hoc adjustments mean that the authentic record of volunteer involvement is often lost in translation.

et al. (1997) tackles the classic problem of employee scheduling by introducing a model that efficiently organizes factory staff schedules, taking into consideration costs, service quality, contractual obligations, and physical limitations. In contrast, Villarreal et al. (2015) offers a more contemporary perspective, developing a mathematical framework to align workforce numbers with demand. This model cleverly addresses the challenges of meeting time-sensitive demand, catering to individual employee characteristics, and operating within the confines of organizational capacity. Ulmer and Savelsbergh (2020) investigates the unpredictable nature of crowdsourced delivery, where maintaining a high standard of service is crucial. To counteract the uncertainty of this approach, some companies prefer a structured delivery system, which ensures a level of predictability. Ulmer and Savelsbergh (2020) provides a strategy for such delivery systems that reduces uncertainty and guarantees exceptional service. While the current body of research yields significant understanding and bears resemblance to our study, this paper forges a novel path by exploring the phenomenon of workers voluntarily donating time and money to the charitable organizations.

Building on the points raised earlier, this paper enriches the existing literature on workforce management by focusing on the unique aspects of volunteer labor. While paid workers and volunteers share some similarities, they diverge significantly when it comes to job satisfaction and motivational underpinnings. The literature identifies a range of motivations, from altruistic to social, that drive individuals to volunteer. For instance, an altruistically inclined volunteer may discontinue their relationship with a charity if they feel their contributions lack meaningful impact, thereby underscoring the risks of overstaffing volunteer events (Clary et al. 1998; Dwiggins-Beeler et al. 2011). In commercial sectors like retail and call centers, labor no-shows also present challenges, affecting customer satisfaction and overall sales (Fisher et al. 2006). Proposed solutions often hinge on employing flexible labor forces, such as part-time or temporary workers (Kesavan et al. 2014; Kamalahmadi et al. 2021). However, an overreliance on such flexible labor can negatively affect performance (Kesavan et al. 2014). Recent field experiments suggest that stable scheduling can both reduce labor costs and boost sales, primarily by enhancing employee effort and decreasing turnout uncertainty (Kesavan et al. 2022).

In volunteer settings, however, such flexible labor policies may be less effective. Unlike paid workers, who are motivated by monetary rewards, volunteers are driven by a desire for meaningful service; they contribute their time to gain an *experience to serve* (Dwiggins-Beeler et al. 2011). As creators of social good, they are less inclined to be available for *on-call* shifts and may feel undervalued if treated merely as backup labor. Moreover, practitioners highlight that the operational context of volunteer events differs substantially from commercial settings. Typically organized as short shifts, labor shortages in volunteer events often become apparent only shortly before their commencement, making it impractical to summon backup volunteers. Additional distinctions arise in organizational objectives. In commercial settings, the objective of workforce management is to cover the demand while minimizing labor costs (Mason et al. 1998), while there is no such labor costs for volunteers (Sampson 2006). Furthermore, unlike in commercial settings where employees have little say in their schedules, volunteers exercise significant control over their availability, adding another layer of complexity to effective workforce management (Sampson 2006).

This paper positions itself within the burgeoning field of research focused on volunteer labor staffing, yet distinguishes itself from the prevailing approaches. For instance, Gordon and Erkut (2004) employ integer programming to construct a scheduling model but overlook the cost implications of volunteer shortages. Sampson (2006) address this gap by using goal programming to minimize a composite cost function that includes labor shortage, over-utilization, and volunteer-task mismatch. However, their model primarily distinguishes between volunteer and commercial labor in terms of cost structures. Building on this, Falasca and Zobel (2012) extend the framework proposed by Sampson (2006), introducing a multi-objective optimization model aimed at assigning humanitarian volunteers to various tasks across multiple locations, while acknowledging that labor costs are not negligible. In the realm of humanitarian relief, Lassiter et al. (2015) employ a robust optimization approach to account for task demand uncertainty, offering a dynamic and flexible framework for volunteer allocation. Further contributing to this discourse, Ata et al. (2019) explore volunteer staffing under conditions of supply and demand uncertainty. Meanwhile, Urrea et al. (2019) examine the influence of volunteer experience on performance outcomes, demonstrating that a homogenous group of volunteers based on experience level performs more effectively than a mixed group. Our study diverges from this existing body of work by introducing a unique set of considerations, offering fresh perspectives and novel solutions for volunteer labor staffing. This paper specifically proposes a model wherein donations and labor are amalgamated within a volunteer workforce.

First, we substantiate our core assumptions through a multi-faceted validation process that encompasses a case study of a prototypical charity's volunteer operations, expert interviews from a renowned nonprofit consulting firm, and an exhaustive review of scholarly works in both social psychology and behavioral economics. Second, we introduce an optimization model that adeptly balances the costs associated with both understaffing and overstaffing, while establishing an explicit linkage between an individual's time commitment and monetary donations. This duality is seamlessly integrated into the charity's overarching objectives. Furthermore, our analytical framework offers volunteer managers straightforward, yet insightful, guidance that can be effortlessly implemented through commonplace tools like Excel spreadsheets for event scheduling. To enrich the practical applicability of our model, we delve into the nuanced heterogeneity of volunteers and account for all pertinent operational constraints. From this, we derive easy-to-interpret decision tree models tailored for various volunteer tasks, thereby offering managers a granular yet comprehensive toolkit for volunteer scheduling.

2. Model

Our central question seeks to determine the optimal number of volunteers of each type to schedule, with the objective of maximizing the total expected utility derived from both time and monetary contributions. In Section 2.1, we initially model the influence of the volunteer composition, denoted by $\mathbf{x}$, on task completion. Subsequently, in Section 2.2, we examine its implications for prospective monetary donations. Lastly, in Section 2.3, we articulate the volunteer staffing conundrum.

Suppose the charity invites $\mathbf{x} = (x_e, x_f)$ episodic and formal volunteers for the volunteering opportunity. Due to their longstanding affiliations with the charity, all scheduled formal volunteers are reliably committed to attending. Conversely, the attendance of episodic volunteers is subject to variability. We capture this stochasticity by introducing a non-negative random variable H, which we designate as the "turnout proportion." We assume that H follows a uniform distribution for two principal reasons. First, within a given interval, the uniform distribution represents the maximum entropy distribution, implying that all potential turnout proportions are equally likely. This assumption is aligned with the operational realities of many charities, which frequently encounter difficulties in managing volunteer data and forecasting turnout rates. Second, the uniform distribution is advantageous for creating parsimonious models with minimal parameters, thereby enhancing the model's interpretability and facilitating insight generation.[4] Although we refer to H as a proportion, it can generally attain a value greater than 1. For the sake of simplicity, we use $\mathbb{E}[\cdot]$ to denote the expectation of random variable H and let h_ℓ and h_u denote the lower and upper bounds of H, respectively. Hence, the number of volunteers who show up to the event is (Hx_e, x_f) with $H \sim U(h_\ell, h_u)$.

2.1. Effect on work completion

Let λ denote the number of volunteer hours required to complete the assigned tasks. Formal volunteers are inherently more efficient than their episodic counterparts. Therefore, without loss of generality, we assume that an episodic volunteer contributes the equivalent of one "volunteer hour," while a formal volunteer contributes θ volunteer hours ($\theta \geq 1$). (If episodic and formal volunteers contribute a and θa volunteer hours, respectively, we can scale demand λ by a.) We also assume that volunteer managers can estimate the value of θ based on their knowledge of volunteering events and interactions with volunteers. Hence, the total labor hours available for the job is $v = Hx_e + \theta x_f$. Since volunteer turnout is random, Hx_e and v are stochastic quantities. The charity management prefers to ensure a sufficient workforce to complete the assigned tasks. Consequently, it gains a per-unit operational benefit $w > 0$ for each unit of work completed, $\min(\lambda, v)$; volunteer tasks that are completed yield a total operational benefit $w \min(\lambda, v)$. The charity incurs an understaffing cost if there

[4] One may choose to apply a binomial distribution to model the turnout uncertainty. While this approach is easy to implement in practice, it becomes infeasible if the average turnout exceeds one. Additionally, the binomial distribution assumes the independence of each Bernoulli trial. However, the arrivals of episodic volunteers are often correlated. Our approach accounts for this correlation by modeling the total number of people who arrive.

is any unfinished volunteer work. A per-unit penalty cost $\tau > 0$ is applied for each unit of unfinished work $(\lambda - v)^+$. Beyond the direct costs and benefits associated with staffing levels, the charity also faces the nuanced issue of overstaffing. An excess of volunteers can undermine individual perceptions of meaningful contribution, leading to diminished engagement with the organization (Smith 1998). We model this as a per-unit overstaffing cost $\gamma > 0$ incurred for any idle volunteer hours, $(v - \lambda)^+$. Accordingly, the charity's total labor gain is

$$L^x(Hx_e, x_f) := w \min(\lambda, Hx_e + \theta x_f) - \tau(\lambda - Hx_e - \theta x_f)^+ - \gamma(Hx_e + \theta x_f - \lambda)^+. \quad (1)$$

We can define a new parameter $\beta := \tau + w$ to be the total labor shortage cost, and rewrite (1) as

$$L^x(Hx_e, x_f) = (w - \beta)\lambda + \beta(Hx_e + \theta x_f) - (\beta + \gamma)(Hx_e + \theta x_f - \lambda)^+. \quad (2)$$

2.2. Effect on monetary donations

Let $d_e > 0$ denote the average increase in monetary donations attributable to the presence of an episodic volunteer. For formal volunteers, the average donation gain consists of two components: one driven by enhanced group identity, and the other influenced by unobserved factors such as altruism or the *warm glow* effect (Andreoni 1990). Specifically, $d_f > 0$ represents the donation gain attributable to these unobserved factors, while $d'_f u(Hx_e, x_f)$ quantifies the donation increase resulting from the formal volunteer's sense of identification with the volunteer group. Here, $u(Hx_e, x_f)$ signifies the utility derived by formal volunteers in a group consisting of Hx_e episodic and x_f formal volunteers. Invoking a slight abuse of notation, we can express u as a univariate function of the transformed variable $\rho := Hx_e/x_f$, which represents the ratio of episodic-to-formal volunteers. This measure finds a parallel in the work of Kesavan et al. (2014), who also quantify labor mix using the ratio of part-time to full-time laborers. While the presence of more formal volunteers strengthens the group identity, the marginal utility from additional formal volunteers exhibits diminishing returns. For instance, once a formal volunteer has established friendships through the volunteer event, the inclusion of subsequent formal volunteers may exert a diminished impact on the overall group identity. Therefore, we assume that u is non-negative and convex decreasing in ρ. For example, u can take an exponential form, $u(\rho) = e^{\alpha - \rho}$, a quadratic form, $u(\rho) = (c - a\rho^2 + b\rho)^+$, an absolute value function form, $u(\rho) = (\alpha - |b - \rho|)^+$, or a linear form, $u(\rho) = (\alpha - \rho)^+$. If (Hx_e, x_f) is the number of volunteers on the day of the job, the charity gains a total benefit associated with monetary donations equal to

$$M^x(Hx_e, x_f) := d_e Hx_e + d_f x_f + d'_f x_f \left(\alpha - \frac{Hx_e}{x_f}\right)^+. \quad (3)$$

2.3. Volunteer staffing decision problem

Let $\mathcal{X}$ represent the set of feasible staffing decisions. Suppose that the volunteer manager is only concerned with the work completed, the staffing decision problem is $\max_{\mathbf{x}\in\mathcal{X}} \mathbb{E}[L^x(Hx_e, x_f)]$. Alternatively, if the manager is concerned with how the staffing decision affects both the current work completion and the future monetary donations, then the staffing decision problem is $\max_{\mathbf{x}\in\mathcal{X}} \mathbb{E}[J^x(Hx_e, x_f)]$, where J^x is the joint objective $J^x(Hx_e, x_f) := L^x(Hx_e, x_f) + M^x(Hx_e, x_f)$. The novelty of the utility function $\mathbb{E}_p[J^x(Hx_e, x_f)]$ is that it compares the trade-off between (i) individuals' time and monetary donations, and (ii) labor shortage and surplus cost. To the best of our knowledge, there exists no analytical research that explores the trade-off between individual time and monetary donations within the context of workforce management. Given the inherent unpredictability of volunteer turnout, the manager possesses only limited information concerning the attendance ratio of episodic volunteers. Specifically, we postulate that the manager can reliably estimate only two parameters, h_ℓ and h_u, based on their wealth of experience. (In Section 5, we further enrich our model by introducing a distribution-free method that allows volunteer managers to leverage existing charity volunteer data while accommodating the inherent ambiguities present within that data.)

We refer to volunteer managers' decision as the *volunteer staffing* (VS) problem. If the volunteer manager is only concerned with work completion, then she can solve the following VS variant:

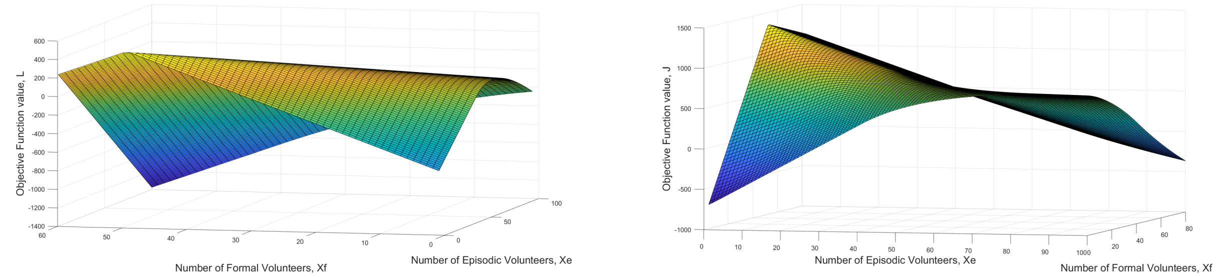
$$L^* := \max_{\mathbf{x}\in\mathcal{X}} L(\mathbf{x}) := \max_{\mathbf{x}\in\mathcal{X}} \mathbb{E}[L^x(Hx_e, x_f)]. \tag{VS-L}$$

We denote this as the VS-L problem, where "L" signifies a labor-oriented objective. Let $\mathbf{x}^*$ be the solution to Equation (VS-L). If the charity proceeds with $\mathbf{x}^*$ volunteers, the expected labor benefit will be denoted by L^*. Should the volunteer manager aim to optimize both task completion and future monetary donations, they can address the following variant of the VS problem:

$$J^* := \max_{\mathbf{x}\in\mathcal{X}} J(\mathbf{x}) := \max_{\mathbf{x}\in\mathcal{X}} \mathbb{E}[L^x(Hx_e, x_f) + M^x(Hx_e, x_f)]. \tag{VS-J}$$

We refer to this problem as VS-J, where "J" refers to a joint objective of labor gain and monetary donations. Suppose $\mathbf{x}^*$ is the solution to (VS-J). If the charity accepts $\mathbf{x}^*$ volunteers, then the charity's expected labor joint benefit will be J^*. Program managers favor maintaining a minimum number of formal volunteers, leveraging their expertise and efficiency to facilitate the volunteer task. Accordingly, we impose a lower bound constraint $x_f \geq x_{f,\ell}$. Additionally, given that the charity operates with a limited pool of formal volunteers, there is an upper limit to their availability for any given task. We model this limitation through the constraint $x_f \leq x_{f,u}$. We also incorporate an upper bound on the number of episodic volunteers, denoted by $x_{e,u}$. Therefore, the feasible set is defined as:

$$\mathcal{X} := \left\{ \mathbf{x} = (x_e, x_f) \in \mathbb{R}^+ \times \mathbb{R}^+ : \begin{array}{l} x_e \leq x_{e,u} \\ x_f \in [x_{f,\ell}, x_{f,u}] \end{array} \right\}. \tag{4}$$



(a) Function $L(\mathbf{x})$ (b) Function $J(\mathbf{x})$

Figure 1 **The functions $L(\mathbf{x})$ and $J(\mathbf{x})$ are plotted against different values of $\mathbf{x} = (x_e, x_f)$. Note the non-concavity of J. For these plots, we set $\lambda = 50, w = \$10, h_\ell = 0.1, h_u = 1.2, \theta = \alpha = 1, \gamma = \beta = 25, d_e = \$15, d_f = \$2.5, d'_f = \$15$.**

3. Optimal staffing decisions under VS

In this section, we derive the optimal staffing decisions under the variants of the VS problem. By abuse of notation, we let $\mathbf{x}^* = (x_e^*, x_f^*)$ refer to the optimal solution where the model (L or J) is clear from the context.

3.1. Model VS-L

We analyze model VS-L whose objective function is $L(\mathbf{x}) := \mathbb{E}[L^x(Hx_e, x_f)]$. From (2), $\mathbb{E}[L^x(Hx_e, x_f)]$ is a newsvendor objective with random yield, $Hx_e + \theta x_f$, and known demand, λ. So, $\mathbb{E}[L^x(Hx_e, x_f)]$ is jointly *concave* in $\mathbf{x}$. The concavity of L is illustrated in Figure 1a. Theorem 1 provides a formal articulation of the optimal solution to Equation (VS-L). The proof is elaborated in Section A.

Theorem 1. *The solution to* (VS-L) *is* $x_f^* = \min\{x_{f,u}, \frac{\lambda}{\theta}\}$ *and* $x_e^* = \frac{(\lambda - \theta x_f^*)\sqrt{\beta+\gamma}}{\sqrt{h_\ell^2\beta + h_u^2\gamma}}$. *The optimal value is* $L_u^* = w\lambda + \frac{(\lambda-\theta x_f)(\beta h_\ell + h_u\gamma - \sqrt{\beta+\gamma}\sqrt{h_\ell^2\beta + h_u^2\gamma})}{h_u - h_\ell}$.

Recall that each formal volunteer contributes θ volunteer hours. Consequently, $\bar{x}_f := \min\{x_{f,u}, \frac{\lambda}{\theta}\}$ represents the maximum number of formal volunteers that can contribute to fulfilling the requisite volunteer hours λ. According to this theorem, when the charity's objective is solely focused on work completion, the optimal strategy is to accept the maximum number of formal volunteers (i.e., $x_f^* = \bar{x}_f$). This result is intuitive because formal volunteers are preferred over episodic volunteers due to their experience at performing the task ($\theta \geq 1$) and their reliability at turning up.

3.2. Model VS-J

Unlike (VS-L) which is a convex optimization problem, the robust staffing problem with a joint objective (VS-J) is generally non-convex. Hence, we generally cannot use efficient convex optimization techniques in solving (VS-J). For analytical tractability, we let $u(\rho) = (\alpha - \rho)^+$, where $\rho = Hx_e/x_f$ is the ratio of episodic-to-formal volunteers. If $\rho > \alpha$, a formal volunteer does not identify with the group since there are too few formal volunteers; hence, their monetary donation is not improved by the group identity. Even in the linear case, the objective function $J(\mathbf{x})$ is neither concave nor convex (see Figure 1b). In what follows, we will first study (VS-J) by deriving the closed-form solution in the case when the constraints (i.e., bounds on the volunteer types) are non-binding. Under binding constraints, the model becomes significantly more complex because the solution depends on which sub-scenario the binding constraints fall into. Therefore, we will delve into computational methods for handling binding cases in Section 5.

This closed-form solution offers significant pragmatic advantages to charitable organizations, as it can be seamlessly integrated into the scheduling procedures employed by volunteer managers. For instance, such a solution could be programmed into an Excel spreadsheet, affording straightforward guidelines for managers concerning the requisite numbers of episodic and formal volunteers for planned events. Furthermore, in the context of relaxed bound constraints (i.e., under the assumption of a sufficient volunteer pool in the long term), this solution provides invaluable insights for program managers regarding the optimal types of volunteers to recruit, tailored to their anticipated events.

If the group composition is (Hx_e, x_f), then under a linear utility, the combined labor and monetary donation gain, $J^x(Hx_e, x_f) := L^x(Hx_e, x_f) + M^x(Hx_e, x_f)$, is:

$$(w-\beta)\lambda + (d_e+\beta)Hx_e + (d_f+\beta\theta)x_f - (\beta+\gamma)(Hx_e+\theta x_f-\lambda)^+ + d_f'\,(\alpha x_f - Hx_e)^+. \qquad (5)$$

Note that $J^x(Hx_e, x_f)$ is neither concave nor convex in $\mathbf{x} = (x_e, x_f)$, a property that carries over to $J(\mathbf{x})$. Hence, we cannot solve (VS-J) using standard convex optimization techniques. However, $J^x(Hx_e, x_f)$ is a piecewise-linear function in H with two breakpoints, $h_0 := \frac{\lambda-\theta x_f}{x_e}$ and $h_f := \frac{\alpha x_f}{x_e}$. The two breakpoints h_0 and h_f represent the labor cost threshold for the newsvendor and the threshold for additional donations from formal volunteers, respectively. This property makes model (VS-J) suitable for closed-form analysis for each possible sub-region. Armed with the closed-form expression for $J(\mathbf{x})$, we are now poised to derive the solution to (VS-J). A challenge is that $J(\mathbf{x})$ is neither concave nor convex (see Figure 1b). Consequently, solving (VS-J) entails evaluating $J(\mathbf{x})$ across six subdomains within the feasible set $\mathcal{X}$. It is crucial to note that the optimal solution is contingent upon the boundary values $x_{f,\ell}$, $x_{f,u}$ and $x_{e,u}$ that delineate the feasible set $\mathcal{X}$. For the sake of model parsimony, we focus on deriving the solution in the scenario where these bounds impose minimal constraints; specifically, when $x_{f,\ell} = 0$ and both $x_{f,u}$ and $x_{e,u}$ are sufficiently large.

First, let us examine the situation wherein volunteering has a significant impact on subsequent donations. We differentiate between two particular scenarios: (1) the maximal monetary gain accrued from formal volunteers surpasses the overstaffing costs, mathematically expressed as $d_f + \alpha d_f' > \theta\gamma$, and (2) the average monetary donation from episodic volunteers exceeds the costs incurred due to the uncertainty of their turnout, formulated as $d_e > \gamma$. It is imperative to note that the donation from a formal volunteer diminishes with the addition of more episodic volunteers to the same team; hence

$d_f + \alpha d'_f$ represents the maximal donation potential in a purely formal volunteer team. Should an episodic volunteer make an appearance when volunteer staffing is already adequate, the organization faces an overstaffing cost denoted by γ. However, the charity may still reap a net positive benefit if the expected average donation from episodic volunteers manages to outweigh the overstaffing cost.

THEOREM 2. *Let $x_{f,\ell} = 0$ and both $x_{f,u}$ and $x_{e,u}$ be sufficiently large. We have the following two cases:*

(a) When $d_f + \alpha d'_f \geq \gamma\theta$, then $x_f^* = x_{f,u}$ and $x_e^* = 0$ if and only if d_e is sufficiently small. Moreover, there exists sufficiently large values of d_e and d_f such that $x_f^* = x_{f,u}$ and $x_e^* = x_{e,u}$.

(b) When $d_e > \gamma$, then $x_e^* = x_{e,u}$ if and only if d_f is sufficiently small. Moreover, there exists sufficiently large values of d_e and d_f such that $x_f^* = x_{f,u}$ and $x_e^* = x_{e,u}$.

Theorem 2 indicates that when volunteering has a significant effect on at least one type of volunteer's donations, the charity can still benefit from inviting an excess number of volunteers. Case (a) holds when $d_f + \alpha d'_f \geq \gamma\theta$. This is when the maximum donation amount in a team of only formal volunteers ($d_f + \alpha d'_f$) exceeds the cost of overstaffing ($\gamma\theta$). In this case, the charity has an incentive to invite the maximum number of formal volunteers, $x_f^* = x_{f,u}$. Moreover, when the episodic volunteers' donation is small, the charity will not benefit from inviting any episodic volunteers because their presence reduces the satisfaction of formal volunteers. Case (b) refers to when the episodic volunteers' average donation, d_e, covers their overstaffing cost. In this scenario, the charity can choose a team composed solely of episodic volunteers. Although formal volunteers are more reliable and may make additional donations if their presence is large enough to form a cohesive group relationship, episodic volunteers can be the sole source of labor supply if their average donation covers the labor cost caused by their random turnout. Lastly, Theorem 2 also states that if the average donation of volunteers is sufficiently large for both types, then it is reasonable to team formal and episodic volunteers together, even though this leads to a less satisfactory environment for formal volunteers. Hence, despite the downside of episodic volunteers as pure labor suppliers, it is desirable for the charity to benefit from this group of volunteers.

While Cases (a) and (b) are realistic for some tasks (e.g., fundraising drives), they may not hold for all volunteering tasks in a typical charity. Hence, we consider the cases where $d_f + \alpha d'_f < \gamma\theta$ and $d_e < \gamma$. In these conditions, the effect of volunteering on donations is moderate, and so the charity should carefully balance the cost and benefit of adding a new volunteer to the group. We define the notation

$$\nu := \frac{h_\ell \beta + h_u \gamma - \sqrt{\beta + \gamma}\sqrt{d_e(h_\ell^2 - h_u^2) + h_\ell^2 \beta + h_u^2 \gamma}}{h_u - h_\ell},$$

Note that ν is composed of terms relating to only the episodic volunteer donation parameter, d_e, turnout statistics, h_ℓ and h_u, and labor cost parameters, β, and γ. We can interpret ν as the marginal value of an episodic volunteer to the charity (that will be further elaborated in Section 4.1). Therefore, ν factors into the decision of whether SVdP prefers formal volunteers or episodic volunteers, as seen in the following theorem.

THEOREM 3. *Suppose* $x_{f,\ell}=0$ *and both* $x_{f,u}$ *and* $x_{e,u}$ *are sufficiently large. If* $d_f+\alpha d'_f<\gamma\theta$ *and* $d_e<\gamma$*, then a solution to* (VS-J) *is:*

(a) *If* $d_f+\alpha d'_f\le\theta\nu$*, then* $x_f^*=0$ *and* $x_e^*=\frac{(\lambda-\theta x_f^*)\sqrt{\beta+\gamma}}{\sqrt{d_e(h_\ell^2-h_u^2)+h_\ell^2\beta+h_u^2\gamma}}$*. The optimal value is* $J^*=w\lambda+\frac{\lambda(h_\ell\beta+h_u\gamma-\sqrt{\beta+\gamma}\sqrt{d_e(h_\ell^2-h_u^2)+h_\ell^2\beta+h_u^2\gamma})}{h_u-h_\ell}$.

(b) *Otherwise,* $x_f^*=\lambda/\theta$ *and* $x_e^*=0$*. The optimal value is* $J^*=w\lambda+(d'_f\alpha+d_f)\frac{\lambda}{\theta}$.

Observe that in Case (a) of Theorem 3, the optimal staffing decision is to invite only episodic volunteers. This scenario underlines the fact that although episodic volunteers may be less efficient and reliable compared to formal volunteers, they can be the optimal choice when their expected donation outweighs both the labor loss incurred due to their unpredictable turnout and the maximum possible donation from formal volunteers (i.e., $d_f+\alpha d'_f\le\theta\nu$). This finding stands in contrast to Theorem 1, where formal volunteers are consistently preferred in the labor-only model (VS-L) due to their superior reliability and labor efficiency. The closed-form expressions in Theorems 1 and 3 allow us to understand how x_e^* is affected by the parameters under models (VS-L) and (VS-J). If the charity invites x_f^* formal volunteers, then $\lambda-\theta x_f^*$ is the total work that episodic volunteers need to fill. When the charity is only concerned with work completion, Theorem 1 suggests that $\bar{x}_e=\frac{\sqrt{\beta+\gamma}}{\sqrt{\beta h_\ell^2+\gamma h_u^2}}$ episodic volunteers is optimal, with the expected work they can produce as $\frac{(h_\ell+h_u)\sqrt{\beta+\gamma}}{2\sqrt{\beta h_\ell^2+\gamma h_u^2}}$. Furthermore, $\bar{x}_e$ increases (resp., decreases) when $\beta>\gamma$ (resp., $\beta<\gamma$). In contrast, Theorem 3 suggests that when the charity considers the donation of its volunteers, the charity should invite more than $\bar{x}_e$ due to the potential donations from volunteers. Hence, when episodic volunteers are donors, then the understaffing (overstaffing) cost must be adjusted up (down) by the donation amount.

4. Implications for process improvements

Given the closed-form solutions in Theorem 1 and Theorem 3, we next discuss several process changes that improve a charity's utility.

4.1. When and how can a charity rely on episodic volunteers?

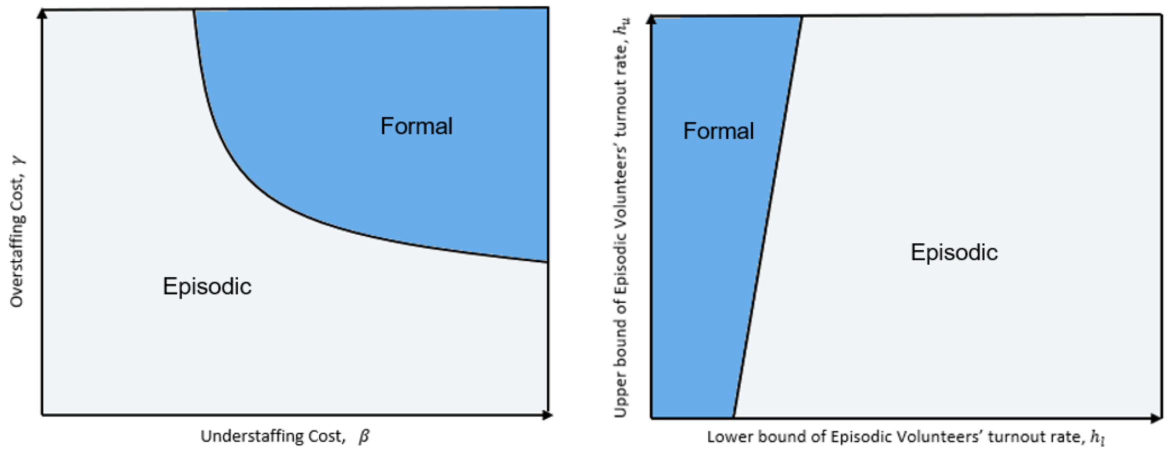
A key question for charities is to understand when they should prefer episodic volunteers as their main workforce. When only considering volunteers' labor value, formal volunteers are clearly the preferred group, due to their reliability and performance. Yet, if charities also consider monetary donations, episodic volunteers could be preferred as the main workforce. In particular, from Theorem 2, we can conclude that charities may prefer episodic volunteers when episodic volunteers' average donation is large enough to cover both the overstaffing cost and the understaffing cost caused by turnout uncertainty.

Further, from Theorem 3 where $d_e < \gamma$ and $d_f + \alpha d'_f < \gamma\theta$, charities should only invite episodic volunteers if the marginal value of episodic volunteers, ν, is larger than the maximum value provided by formal volunteers, $(d_f + \alpha d'_f)/\theta$. Note that the value of episodic volunteers decreases when labor cost (γ or β) increases ($\frac{\partial \nu}{\partial \beta} < 0$ and $\frac{\partial \nu}{\partial \gamma} < 0$). Simply put, although episodic volunteers do bring labor value to complete a task, their net contribution to the charity decreases as labor value becomes greater. When both labor costs (β and γ) are significant, it is optimal to invite only formal volunteers. This can be observed in the dark shaded region of Figure 2a; Although the increase in the monetary donation by episodic volunteers ($d_e = \$15$) is greater than the largest possible increase by formal volunteers ($d_f + \alpha d'_f = \$5$), the benefit does not overcome the high expected cost from the uncertain turnout of episodic volunteers. This observation brings us to the second element that should be considered in this trade-off: the impact of episodic volunteers' turnout uncertainty. We can consider the disparity between h_u and h_ℓ as the proxy of turnout uncertainty, and denote $\epsilon = \frac{h_u}{h_\ell}$ as the degree of uncertainty. Note that $\epsilon \in [1, \infty)$ and equal to 1 when $h_u = h_\ell$. Under model (VS-J), charities should only utilize episodic volunteers when the variance of episodic volunteers' turnout is small enough. Figure 2b demonstrates how the optimal staffing plan is affected by h_u and h_ℓ from Theorem 3. Note that the linear threshold that distinguishes the two policies is $\epsilon = \frac{h_u}{h_\ell} = \frac{(d_f+\alpha d'_f)^2+2(d_f+d'_f\alpha)\beta\theta-(d_e(\beta+\gamma)+\beta\gamma)\theta^2}{(d_f+\alpha d'_f)^2-2(d_f+d'_f\alpha)\gamma\theta+(d_e(\beta+\gamma)-\beta\gamma)\theta^2}$ (equivalent to $d_f + d'_f\alpha = \theta\nu$), which is the smallest disparity between h_u and h_ℓ where the optimal policy is to invite episodic volunteers. Therefore, when the difference between h_u and h_ℓ is larger than this threshold, the optimal policy is to only invite formal volunteers.

We next discuss how ϵ impacts the charity's utility. For (VS-L), the charity is worse off when ϵ increases because $\frac{\partial L^*}{\partial \epsilon} < 0$ in Theorem 1. For model (VS-J), under the conditions of Theorem 3, we can also check that $\frac{\partial J^*}{\partial \epsilon} \leq 0$, where the inequality is strict if $d_e < \gamma$. Hence, when only volunteers' labor value is considered (Theorem 1) or volunteers' labor value is more significant than their donation potential (Theorem 3), the charity is better off reducing the variability of episodic volunteers' turnout by either reducing either the no-show or over-show scenarios. To reduce overstaffing caused by turnout uncertainty, charities may require volunteers to sign up each individual who is planning to participate. Furthermore, research shows that understanding the value of the volunteering task can strengthen volunteers' psychological contract (Vantilborgh et al. 2012) that is likely to reduce the chances of their no-shows. Therefore, charities may consider communicating the importance of the task to volunteers, for example, through emails or text messages, to minimize the likelihood of absenteeism.

4.2. Training programs for formal volunteers

A common practice adopted by charities is to provide additional training programs to their formal volunteers. We analyze whether or not this practice is always beneficial. Consider a charity that is



(a) Cost regions of problem (VS-J) in Theorem 3

(b) Uncertainty set regions of problem (VS-J) in Theorem 3

Figure 2 **Region of costs (β and γ) and uncertainty set (μ and σ) that determine the optimal staffing strategy. In both panels, we set $h_\ell = 0.1, h_u = 1$, $\theta = 1$, $\alpha = 1$, $\beta = 30$, $\gamma = 30$, $d_e = 15$ and $d_f = d'_f = 2.5$.**

only concerned with work completion when making staffing decisions. For example, volunteers at health clinic charities or mental health charities are usually valued for their expertise and experience with the job, not monetary donations. From Theorem 1, we can check that $\frac{\partial L^*}{\partial \theta} \geq 0$, where the inequality is strict if $\theta < \lambda/x_{f,u}$. So, the charity's utility always increases with the efficiency level θ. The intuition behind this is that if formal volunteers are more efficient, fewer episodic volunteers need to be invited, which benefits the charity since the latter group introduces turnout uncertainty. Therefore, the charity should invest in training and developing formal volunteers' skills at completing the tasks. However, once the efficiency level of formal volunteers is sufficiently high (i.e., $\theta \geq \lambda/x_{f,u}$), the charity does not need episodic volunteers to complete the job. Further training will not yield additional benefits to the charity and $\frac{\partial L^*}{\partial \theta} = 0$.

Interestingly, if the charity is concerned with both work completion and monetary donations when staffing jobs, additional training to formal volunteers could *decrease* the charity's utility. When the optimal solution is to only invite formal volunteers, then additional training always decreases the charity's utility since $\frac{\partial J^*}{\partial \theta} < 0$. The intuition is that, as θ increases, the charity will need fewer formal volunteers to complete the job, resulting in lower total donations for the invited volunteers. Similarly, when the optimal solution is a mixture of formal and episodic volunteers, the charity can actually be worse off by training its formal volunteers. The intuition is that, when formal volunteers become more efficient, the charity will require fewer of both types of volunteers. Although this benefits work completion due to having fewer (unreliable) episodic volunteers, this ultimately hurts the charity due to a larger reduction in the charity's future donations.

5. Application: A case study of the Society of St. Vincent de Paul

With 800,000 members in 153 countries across six continents, the Society of St. Vincent de Paul (SVdP) is an international humanitarian organization serving more than 30 million people globally. Their services include feeding, clothing, housing, and healing individuals. With a volunteer-to-staff ratio of 16 to 1, SVdP has nearly 100,000 trained volunteers across 4,400 communities in the U.S. that together provided 12.6 million hours of volunteer services during 2017. Its largest division in the U.S. is located in Phoenix, Arizona, where it serves homeless and low-income families with services such as free medical and dental clinics, food warehouses, transition, and housing. Currently, SVdP Phoenix has about 300 regular employees, over 2,500 active and associate members, and more than 6,000 volunteers. In 2019, SVdP used more than 705,400 volunteer hours and provided 2.6 million meals to people in need. (Figure 3 shows examples of SVdP's operations.)

(a) Pizza Friday (b) Dining Rooms

Figure 3 Examples for the most frequent volunteering tasks

To commence their volunteering at SVdP, individuals can register through the organization's online platform or by phone, selecting their areas of interest and indicating their available time slots based on the location of the volunteer activities. Although most volunteer roles do not necessitate a formal interview or comprehensive vetting process,[5] the organization periodically offers optional orientation sessions to acquaint prospective volunteers with the array of opportunities at hand. Armed with this information, program managers carefully plan upcoming events and determine the number of volunteers needed for each role. Personalized invitations for these pre-arranged tasks are then sent out to volunteers. This well-organized method eliminates the usual complexities related to volunteer-to-task matching that some organizations face. For those volunteers wishing to extend their engagement with SVdP, the process is simple: they just need to contact the volunteer coordinator to update their availability. Before the advent of the COVID-19 pandemic, the system even allowed for

[5]Screening process is different for professional medical volunteers.

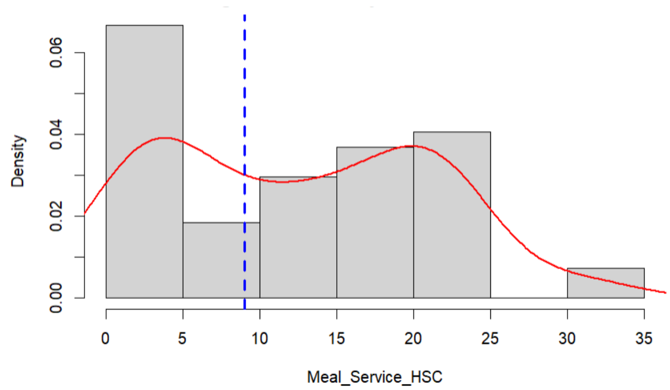


Figure 4 **Episodic volunteer turnout count histogram and density graph on 54 consecutive "Meal Service" events. The red curve line represents the density curve, and the blue dotted line represents the ideal number of volunteers that is nine. Nearly half the time, more than 15 volunteers show up, highlighting a significant over-show issue.**

advance bookings, up to a year before the scheduled volunteer events. Nevertheless, as highlighted in the Introduction, the no-show rate among episodic volunteers averages around 30%. Most absentees either neglect to notify the coordinators or do so too late, making it impossible to find a last-minute replacement. Consequently, understaffing poses a significant challenge, leading to unmet needs and elevated operational costs. For example, when SVdP organizes events like "Pizza Friday" or "Dining Rooms," materials are prepared ahead of time. In the event of a labor shortage, these pre-allocated resources go to waste, and SVdP falls short in providing meals to those it aims to serve. SVdP has also been dealing with the challenge of overstaffing because episodic volunteers often spontaneously invite friends or family to join them without notifying the charity in advance. As a result, management now requires individual volunteer sign-ups and only reveals location assignments a day before events. However, overstaffing issues persist. Figure 4 illustrates this challenge by showing that the number of attendees often far exceeded the ideal team size of 9. In response, SVdP primarily invites formal volunteers, filling remaining spots with episodic ones. (This established approach serves as the "base" policy in our numerical experiments.)

5.1. Lack of data, and distributionally robust optimization

Although the closed-form solutions provided by Equations (VS-L) and (VS-J) offer straightforward policies and managerial insights, charities can further improve their staffing practices if they collect and utilize volunteer data. However, as highlighted in the Introduction, data quality remains a challenge in volunteer management across charities. Prior to March 2018, SVdP's data collection was markedly subpar. Even recent data from 2018 to 2020 reveals substantial issues: nearly 40% of all records contain errors, complicating any reliable analysis. SVdP datasets, for instance, suffer from manual entry errors, such as an incorrect entry of "no-show" or "confirmed" status. The net result is a loss of accurate volunteer records. We, therefore, adopted a *distributionally robust optimization* (DRO) approach, which obviates the need to specify a distribution for the unpredictable volunteer no-shows. This approach is well-suited for the inherently unpredictable nature of volunteer attendance. Note that even if SVdP had perfect data, predicting a volunteer's likelihood to attend remains an uphill battle due to the numerous unknown variables affecting their decision. Further complicating matters is the fact that most variables that could help predict volunteer turnout—such as age, gender, or address—are either self-reported, non-mandatory, or often erroneous. Additionally, episodic volunteers, who make up an average of 84% of SVdP's volunteer base, pose particular challenges for data collection and prediction. Therefore, the volunteer invitation policy is largely shaped by task requirements and generalized historical data rather than precise predictive analytics.

The DRO approach was popularized by Scarf (1958) for the classical newsvendor model, where the objective was to choose a solution that maximizes the worst-case expected profit over the set of distributions with mean μ and standard deviation σ. We will adopt a DRO approach for the charity's volunteer management problem, where the optimal volunteer plan is guaranteed to be robust against all possible distributions, as it is optimized based on the worst-case scenario among all distributions. Recall that H represents the random turnout proportion. The DRO approach leverages imperfect historical data and conservatively assumes the first- and second-moment information without specifying the distribution family. Moreover, the simple information required can be supplemented and corrected by volunteer managers who have a good sense of the popularity of the volunteer task and volunteers' reliability. In this way, charities' decisions are protected against unobserved shocks. Let $p:[0,\infty]\mapsto\mathbb{R}^+$ be the probability density function of H. We use $\mathbb{E}_p[\cdot]$ to denote the expectation under distribution p. We let μ and σ denote the mean and standard deviation of H. We also assume that H is non-negative and has bounded support, where h_ℓ and h_u are the lower and upper bounds, respectively ($0\le h_\ell<h_u$). Hence, p must belong to the distribution set:

$$\mathcal{P} := \left\{ p:[h_\ell,h_u]\mapsto\mathbb{R}^+ \;:\; \begin{array}{l} \mathbb{E}_p[1]=1 \\ \mathbb{E}_p[H]=\mu \\ \mathbb{E}_p[H^2]=\mu^2+\sigma^2 \end{array} \right\}. \tag{6}$$

In this context, $\mathcal{P}$ denotes the set of all probability distributions with support within the interval $[h_\ell,h_u]$, characterized by a mean of μ, and a variance of σ^2. Therefore, a DRO solution to the volunteer staffing problem can be formulated as follows:

$$J^*_{DRO} := \max_{\mathbf{x}\in\mathcal{X}} J_{DRO}(\mathbf{x}) := \max_{\mathbf{x}\in\mathcal{X}} \inf_{p\in\mathcal{P}} \mathbb{E}_p[L^x(Hx_e,x_f)+M^x(Hx_e,x_f)]. \tag{DRO-J}$$

Similar to Equation (VS-J), the (DRO-J) is neither convex nor concave. Unlike Equation (VS-J), however, the objective function J^*_{DRO} transforms the problem into a max-min formulation, significantly elevating its computational complexity. In fact, the inner problem of J^*_{DRO} is a piecewise function with up to thirty-two subdomains of $\mathcal{X}$. However, for a fixed $\mathbf{x}$, the value of $J_{DRO}(\mathbf{x})$ can found efficiently using convex optimization techniques since it can be reformulated as a second-order cone program. Hence, we can solve (DRO-J) using grid search where each iteration solves a convex optimization model. A detailed description of the reformulation is deferred to Section C. In our numerical experiments, this distributionally robust approach is referred to as the "DRO" policy.

5.2. Numerical experiments

Four of SVdP's primary volunteer engagements include the Family Evening Meal (FEM), Building Temporary Shelter (BTS), Fundraising (FUN), and Resource Center (REC) initiatives. Each of these events exhibits unique characteristics in terms of labor uncertainty, cost implications, and operational requirements. A nuanced understanding of each event's unique requirements and challenges informs our broader analysis and recommendations for volunteer management at SVdP. The Family Evening Meal (FEM) aims to alleviate hunger by preparing and serving meals in one of SVdP's dining centers. Based on our consultations with volunteer managers at SVdP, FEM incurs substantial costs for both understaffing and overstaffing. While it is imperative to have sufficient volunteers to meet the demand for meals, the space constraints of the dining center limit the number of volunteers that can be accommodated, leaving any surplus volunteers underutilized. Building Temporary Shelter (BTS) is another crucial initiative, particularly given Phoenix's extreme weather conditions. During this event, volunteers collaborate to construct temporary heat relief shelters and offer additional support to residents. Similar to FEM, BTS experiences high understaffing costs; the time-sensitive nature of shelter provision means any delays could negatively impact beneficiaries. However, overstaffing costs are minimal, given the multi-faceted and ongoing nature of the work involved. Thirdly, SVdP hosts specialized Fundraising (FUN) events, which are characterized by low overstaffing costs. Volunteers are dispersed across multiple locations to solicit donations and can operate independently, thus reducing the likelihood of redundancy. Lastly, the Resource Centers (REC) provide a sanctuary for individuals in precarious situations, such as those experiencing homelessness. These centers offer an array of services including showers, clothing, counseling, and referrals. Given the diverse range of tasks and the necessity for adequate volunteer coverage, we posit that REC encounters low overstaffing costs but significant understaffing costs.

The summary statistics for various volunteer tasks are shown in Table 1. We calculated the value of w based on the compensation provided to paid staff for completing each respective task. The understaffing cost, β, incorporates both w and the cost of essential materials. Additionally, we established lower and upper bounds for the number of formal volunteers per event, guided by the range observed in SVdP's data. The overstaffing cost, γ, can be approximated by assessing its influence on future volunteer participation. For instance, if a volunteer perceives an excess of volunteers at an event, diminishing their role, it reduces the likelihood of them returning as a

Table 1 Summary statistics for the sampled tasks

Volunteer task	Instances	λ	μ	σ^2	h_ℓ	h_u	$x_{f,l}$	$x_{f,u}$	w	β	γ
Family Evening Meal (FEM)	93	25	0.85	0.06	0.3	1.2	5	15	20	30	$[10, 20]$
Building Temporary Shelter (BTS)	31	20	0.67	0.05	0.5	1	5	15	30	45	$[5, 10]$
Fundraising (FUN)	19	50	0.65	0.08	0.5	1.2	5	10	20	25	$[5, 10]$
Resource Center (REC)	55	15	0.62	0.04	0.4	1.5	2	5	30	40	$[5, 10]$

volunteer or making a monetary contribution in the future. Rather than utilizing fixed numerical values for overstaffing costs, we consider a range for $\gamma \in [10, 20]$, corresponding to a critical ratio range from 60% to 75%, and then randomly draw 100 instances. Each unique pair of (β, γ) serves as a distinct observation for the computational experiment. Owing to the paucity of available data, we rely on the expertise of SVdP's volunteer managers to estimate the uncertainty surrounding volunteer turnout, represented by μ and σ.

To conduct a numerical assessment of the efficacy of our proposed solution, we first set a benchmark by examining SVdP's conventional volunteer scheduling methodology. Aforementioned, volunteer managers at SVdP give precedence to formal volunteers for all assignments, employing episodic volunteers only to address any shortfall. We term this established heuristic as "Base." As suggested by a reviewer, we introduce an additional heuristic grounded in the classic newsvendor model, which is modified to account for the nuances of donations. This modified approach, referred to as NVD (Newsvendor with Donations), simplifies the VS-J model by retaining the concavity of newsvendor model. Importantly, it also accounts for the potential impact of volunteers' donations. The re-calibrated understaffing cost for episodic volunteers, denoted by β', incorporates three distinct elements: the original labor shortage cost β, the contributions from episodic volunteers d_e, and supplemental donations from formal volunteers ascribed to their sense of group identity d'_f, and $\beta' = \beta + d_e - d'_f$. Put differently, when an episodic volunteer is absent, the charity not only incurs a labor shortage cost but also forgoes the average donation typically contributed by episodic volunteers. This absence may also lead to an uptick in donations from formal volunteers, owing to a stronger sense of group identity when the team is composed primarily of formal volunteers. On the flip side, the redefined overstaffing cost, denoted by γ', consists of two facets: the original overstaffing cost and a potential reduction in donations from formal volunteers (i.e., $\gamma' = \gamma + d'_f$), attributable to the heightened presence of episodic volunteers. It is worth noting that within the framework of the newsvendor model, the charity will invariably prioritize formal volunteers initially, a strategy congruent with the "Base" policy traditionally employed by SVdP. However, the NVD approach extends this by thoughtfully incorporating the variable of potential donations. Moreover, we make the assumption that the turnout of episodic volunteers is uniformly distributed when employing the NVD policy.

Next, we compare the optimality gaps of these two heuristic policies and our proposed uniform (UNF) as well as DRO policies under four distributions; truncated normal (TN), uniform (UN), u-quadratic (UQ), shifted beta (B). All distributions use the same support, with the TN and B distributions incorporating additional first and second moment information. To provide a benchmark for all policies, we consider a clairvoyant solution assuming that full information about the specific distribution F is known. The utility value of the clairvoyant policy is the upper-bound, denoted as J_F^*. Next, we calculate the expected value using our three established policies $J_{Base}, J_{NVD}, J_{UNF}$ and J_{DRO} and the optimality gap represents the performance of each policy. We consider $d_e = \$4.6$, $d_f = \$2.35$ and $d'_f = \$1.5$.[6] We assume formal volunteers' efficiency level is $\theta = 1.2$ and $\alpha = 1$.

All numerical examples are implemented in Matlab. The inner problems (SOCP problem) are solved with `cvx` package in Matlab. On average, each problem takes 230 seconds to solve on a 1.8 GHz 4-Core Intel Core i7 processor.

We ascertain the average optimality gap between J_F^* and $J_{Base}, J_{NVD}, J_{UNF}$ and J_{DRO} across diverse tasks. For every observation, solutions are generated according to the four distinct policies, thereafter tested under four distributions. Their overarching performance is also evaluated by taking the weighted average of the optimality gap across all four distributions. Table 2 summarizes the mean and standard deviation of the optimality gap of the policies for 16 experimental conditions (combinations of four volunteering tasks and four distributions). The policy bearing the lowest optimality gap for each experimental condition is accentuated (in blue and bold font). Using the Base policy performance as the benchmark, we evaluated the policies by calculating the median difference in optimality gap across 1600 instances (with 100 instances for each experimental condition).[7] We discern a significant enhancement in performance when charities duly account for the donation value of volunteers. The combined labor and donation values increased by a median value of 6.41% upon juxtaposing the Base policy with the UNF policy. This underscores the opportunity cost incurred

[6] We estimate the donations from the data shared by SVdP. First, we categorized individual volunteers into two types: *Formal* and *Episodic* volunteers. Hustinx et al. (2008) specify episodic volunteers as those who has showed up less than once a month. Likewise, volunteer managers in SVdP use participation frequency and volunteering lifetime (duration between first and last volunteering event) as the distinguishing criteria. They consider those who have participated frequently and have volunteered more than six months as formal volunteers. According to volunteering data and given the fact that some formal volunteers "leave" the charity, we adopt both definitions and strictly define those who have participated at least 24 times and volunteering lifetime at least 6 months as formal volunteers. Then, we matched the volunteering data with donations. We find that episodic volunteers on average donate \$18.6 per volunteer event, and formal volunteers on average donate \$15.6 per attendance. To estimate the impact of one's volunteering on her subsequent donations, we use the identified individuals who had registered to volunteer at SVdP but never donated before ($N = 13,511$). Comparing the donation amount between those who registered and showed up and those who registered but did not show up, we found that individuals who completed their volunteering service donated, on average, \$16.4 (SD = 286.8), while those who did not show up donated, on average, \$8.3 (SD = 105.4). This 49.4% loss is statistically significant at $p \leq 0.1$ level. Last, we estimate the donation increase due to volunteering by applying 49.4% to both values. Therefore, episodic volunteers on average donate \$9.2 per volunteering event while formal volunteers on average donate \$7.7 per attendance. To be more conservative, we took another 50% off from the estimated value.

[7] If we exclude the Uniform distribution condition, leaving us with 1200 instances, the median performance improvement is 4.94% for the UNF policy and 10.60% for the DRO policy.

when charities solely reckon the labor value of volunteers and overlook their potential donation values.

Furthermore, the DRO policy demonstrates a median increase of 8.87% in the charity's utility. This strong performance is particularly notable, as the DRO policy consistently exhibits the lowest optimality gap in seven out of the sixteen conditions and secures the overall lowest optimality gap in two out of the four volunteering tasks. This underscores the merits of a robustness-centered approach. Additionally, the optimality gap of DRO showcases a narrow range, signifying that the DRO policy stands as a distribution-agnostic strategy that assures commendable performance under a variety of distributions. A more vivid illustration of the robustness value is depicted in Section 5.2, which presents a violin chart of the optimality gap for all volunteering tasks. The skewed values of the DRO policy delineates a clear advantage in circumventing severe operational deficits.

Thirdly, the optimality gap for the informed newsvendor model (NVD) can be significant for the FUN and REC tasks. This occurrence stems from the newsvendor structure's balance between understaffing and overstaffing costs, operating under the assumption that an excess of volunteers is undesirable. Nonetheless, when volunteers' donations are taken into account, it may transpire as optimal for charities to extend invitations to a greater number of volunteers, even if it engenders overage situations (e.g., FUN events). Despite this, the NVD policy can outperform other policies in the FEM task, wherein both understaffing and overstaffing costs far exceed the donation values.

Lastly, we compare the staffing choices of Base, NVD, UNF, and DRO. The average count of formal and episodic volunteers enlisted by each policy is illustrated in Table 2. A notable augmentation in the number of episodic volunteers is observed when accounting for both labor and donations (UNF, DRO), as compared to considering solely labor (Base), reaffirming our inference that a heightened reliance on episodic volunteers is prudent when contemplating volunteers' donations.

Each of the four evaluated policies has its own array of merits and challenges. For instance, UNF exhibited strong performance in certain volunteering events where the underlying distribution followed either a U-Quadratic or Truncated Normal distribution. We attribute this success to the symmetric nature of these two distributions. In contrast, the DRO policy excelled particularly when the underlying distribution became more irregular, as seen in the case of the Beta distribution. The NVD heuristic also demonstrated noteworthy performance in several instances, especially when formal volunteers were preferred or when labor costs were higher than donations. However, NVD's performance declined significantly when the episodic volunteers became more desirable (e.g., when labor costs were lower). In summary, volunteer managers may choose to implement the UNF policy when they observe a consistent and stable volunteer turnout. Conversely, if volunteer turnout becomes volatile, opting for the DRO policy may be more prudent. Similarly, when labor costs are a primary concern for the charity, the NVD policy emerges as a dependable choice, offering reliable performance under various circumstances.

Table 2 Optimality Gap on 100 instances under four different distributions. The value in parenthesis is the standard deviation. The "Overall" value is the weighted average optimality gap from four distributions. All values are in percentage except for average x_e and x_f, which are in absolute numbers. For the conditions under uniform distribution, we highlighted the second best solution since $J_F^* = J_{UNF}$.

Volunteering Task	Distribution and Policy	Base	NVD	UNF	DRO
FEM	TN	7.79 (2.93)	**0.01 (0.00)**	4.88 (2.17)	1.92 (0.19)
	UN	22.2 (5.98)	**3.11 (0.76)**	0.00 (0.00)	12.04 (1.67)
	UQ	25.83 (19.57)	19.39 (11.21)	**17.72 (19.80)**	21.93 (14.84)
	BETA	3.32 (2.73)	0.98 (0.58)	8.21 (3.12)	**0.22 (0.39)**
	Overall	12.27 (6.05)	**3.05 (1.67)**	5.64 (2.06)	6.23 (2.92)
	Average x_f	15.0	15.0	15.0	15.0
	Average x_e	8.0	9.7	10.8	8.8
BTS	TN	4.20 (3.35)	**0.88 (2.85)**	1.44 (1.65)	3.56 (4.68)
	UN	6.62 (6.11)	3.93 (5.66)	0.00 (0.00)	**1.52 (2.05)**
	UQ	12.45 (5.45)	7.19 (4.78)	4.30 (3.27)	**1.02 (2.14)**
	BETA	26.58 (4.53)	19.38 (3.51)	23.64 (7.71)	**0.02 (0.01)**
	Overall	12.69 (4.87)	8.04 (4.19)	7.68 (2.74)	**1.49 (2.15)**
	Average x_f	15.0	15.0	9.4	7.9
	Average x_e	3.0	3.3	16.3	21.0
FUN	TN	4.75 (2.46)	13.62 (0.14)	**0.25 (0.37)**	5.36 (3.90)
	UN	**4.19 (2.46)**	9.10 (1.78)	0.00 (0.00)	5.04 (4.45)
	UQ	12.13 (3.43)	22.77 (0.75)	3.06 (2.06)	**2.25 (2.24)**
	BETA	52.6 (0.18)	70.05 (4.95)	34.21 (14.25)	**0.05 (0.01)**
	Overall	19.46 (1.71)	30.11 (1.02)	10.19 (4.64)	**3.05 (2.55)**
	Average x_f	10.0	10.0	5.0	5.0
	Average x_e	58.5	52.4	75.9	88.0
REC	TN	8.68 (2.90)	30.19 (1.70)	**3.62 (2.92)**	4.83 (3.59)
	UN	**2.58 (2.79)**	8.62 (3.02)	0.00 (0.00)	11.85 (10.26)
	UQ	8.14 (4.31)	20.82 (3.15)	**3.69 (1.73)**	7.94 (8.19)
	BETA	21.01 (3.11)	44.24 (1.54)	13.35 (6.32)	**0.10 (0.10)**
	Overall	10.37 (3.14)	26.53 (0.49)	**5.38 (2.80)**	6.18 (5.51)
	Average x_f	5.0	5.0	3.1	2.0
	Average x_e	14.25	11.2	20.7	31.5

These policies also vary in terms of their practical ease of implementation. For instance, while the DRO policy stands out with superior overall performance, charities might grapple with hurdles related to tracking and collecting volunteers' attendance data, potentially stemming from relationship management intricacies or inadequate data infrastructure. In light of such challenges, the UNF policy surfaces as a compelling alternative, situating itself as a robust second-best choice. On the flip side, the NVD policy, with its facile implementation and intelligibility, presents a particularly appealing option for charities. This policy seems especially apt for volunteering tasks mirroring the dynamics of the FEM tasks, often laden with significant understaffing and overstaffing costs. At the core, the efficacy of any chosen policy is firmly anchored on the bedrock of reliable data availability and a nuanced understanding of volunteer behaviors, highlighting the importance of informed decision-making in enhancing operational effectiveness within charitable domains.

5.3. Applications and Insights

We have engineered a user-friendly, Excel-based decision support tool that calculates the staffing decisions under the policies used in our numerical experiments. This intuitive tool, accompanied by preset examples, necessitates no advanced solver, thereby making it accessible for volunteer

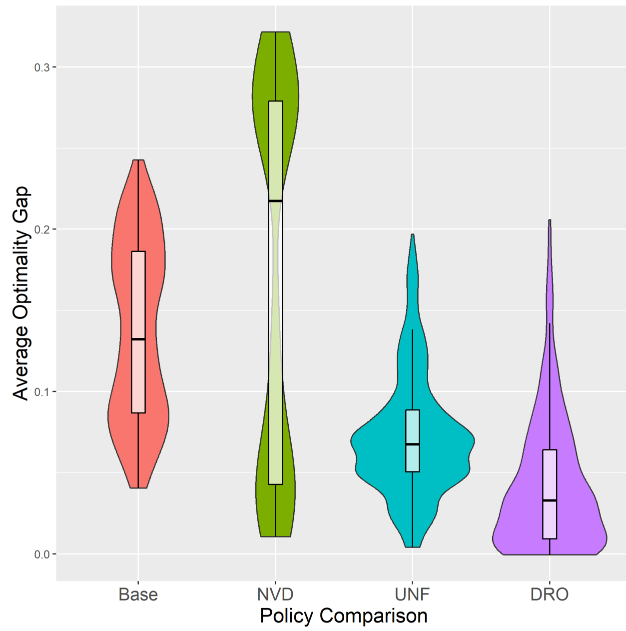


Figure 5 Performance evaluation of four policies for all volunteering tasks under four distributions. The figure combines box-plot and violin chart to demonstrate both the median and distribution of the plots. As expected, DRO policy is shorter in terms of optimality gap.

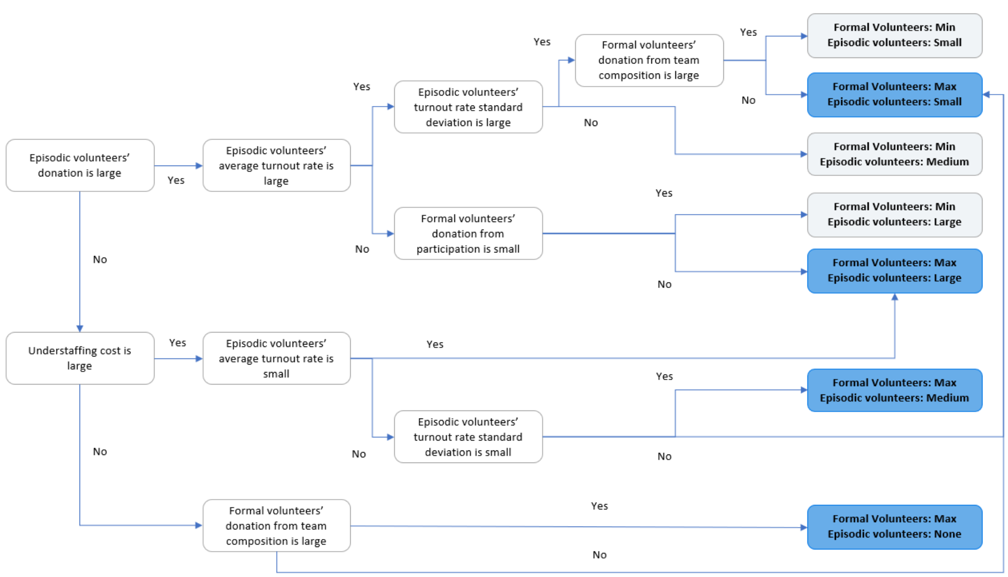


Figure 6 Decision tree: A general simplified process to determine workforce configuration

managers. It offers a visual conduit through which managers can seamlessly perceive how the optimal policy morphs in response to varied parameters, and juxtapose multiple policies under specific parameter sets. This interactive tool is designed not only to streamline decision-making but also to illuminate the tangible impact of different policy choices, aiding managers in making informed, strategic decisions in volunteer engagement and resource allocation.

Although the DRO solution showcased its robustness against distribution uncertainty, the computational method could be difficult to incorporate into a charity's existing processes since it requires a convex optimization solver. Instead, we provided the charity with an interpretable decision model (decision tree) that has been trained from the solution of 100,000 randomly generated instances[8] of constrained (DRO-J). The decision tree allows us to develop insights into how the parameters affect the optimal solution. It is worth noting that this approach can also be tuned for a specific volunteering task, and hence generated prescriptive insights for charities. We removed instances where the optimal policy is to invite overwhelming amount of volunteers (i.e., $x_e^* = x_{e,u}$) because this only represents less common cases (e.g., $d_e > \gamma$) and the insight is simple to conclude. Figure 6 presents the decision tree, where the dark (light) shaded leaf nodes correspond to optimal solutions where the formal volunteers are equal to the upper (lower) bound.

As shown in Figure 6, the factor that best splits the data is the magnitude of d_e, the effect of volunteering on episodic volunteers' donation. When d_e is small, it is optimal to admit the maximum formal volunteers. The secondary factor is $d_f + \alpha d'_f$, the maximum donation from a formal volunteer. When the $d_f + \alpha d'_f$ is large, the optimal policy is to invite the maximum number of formal volunteers. Otherwise, the optimal policy is to invite a minimum number of formal volunteers. Finally, the factors that determine the optimal number of episodic volunteers are the average (μ) and standard deviation (σ) of episodic volunteer turnout. When μ is low, it is optimal to select the maximum number of episodic volunteers. On the other hand, when μ is large, it is optimal to invite the minimum or moderate number of episodic volunteers, depending on σ. Specifically, a more extensive σ means a lower number of episodic volunteers.

Another interesting insight is when the constraints are binding and $x_{f,u} < \lambda/\theta$. Under some conditions (see Figure 6), the optimal number of formal volunteers is $x_f^* = x_{f,u}$, so training these volunteers does not impact the number of formal volunteers but decreases the episodic volunteers needed. If d_e far exceeds d'_f, the fewer episodic volunteers will have a negative net effect on the charity's utility; Although the formal volunteers will increase their donation due to the decrease in ρ (episodic-to-formal volunteers ratio), it is not enough to compensate for the loss of monetary donation from the fewer episodic volunteers. Accordingly, more experienced formal volunteers could crowd out other volunteers' participation, thus possibly decreasing the total monetary donation. This conclusion is also confirmed by SVdP's volunteers survey. For example, a volunteer commented "There is a couple who had volunteered a couple of years before and they are clearly dedicated and generous volunteers. Here's the however: they arrive 90 minutes before the scheduled time and get the dining room set

[8] To generate these instances, we uniformly draw β, γ from $[5, 35]$, d_e from $[0, 35]$, d_f, d'_f from $[0, 25]$, α, θ from $[1, 2]$, h_u (h_ℓ) is drawn from $[0.25, 0.45]$ ($[0.9, 1.1]$), μ from $[0.45, 0.9]$, and σ from $[0, \sigma_u]$ where $\sigma_u = \sqrt{(h_u - \mu)(\mu - h_\ell)}$ is the upper bound by Bhatia-Davis inequality. The lower and upper bound of formal volunteers ($x_{f,l}$ and $x_{f,u}$) are draw uniformly between $[0, \lambda h_\ell / (\alpha h_u + \theta h_\ell)]$ and $[\lambda h_u / (\alpha h_\ell + \theta h_u), \frac{\lambda}{\theta}]$. Last, we set $\lambda = 50$ and $x_{e,u} = \frac{\lambda}{h_\ell}$.

up and many tasks accomplished. While on one hand that is great, on the other, it means there is nothing for others to do who arrive at the published start of the shift. I totally appreciate the kind-hearts of the couple who arrive early, but also think it discourages other volunteers." Since formal volunteers have a longstanding relationship with the charity, they can also gain experience and skills over time if they are assigned to repetitive tasks. Likewise, the charity could spend resources creating challenging and rewarding volunteering tasks (so that formal volunteers' efficiency does not crowd out other volunteers) rather than training volunteers on existing work. Alternatively, charities can position the experienced volunteers to take the leading role to guide other volunteers instead of finishing the jobs by themselves. In addition, they can design tasks such that the total demand is flexible (and so overstaffing cost will likely be small) such that formal volunteers will not crowd out other volunteers' experience. In short, charities should not exclusively concentrate on enhancing the efficiency of their volunteers in completing tasks. Instead, they should consider employing a range of strategies to effectively manage volunteering events. Ultimately, volunteering is not merely an obligation of labor duty but also an opportunity to provide a meaningful charitable giving experience.

6. Conclusion

This paper studies the volunteer staffing problem from a strategic perspective. To develop a nuanced understanding of the actual situation, we collaborated with a large local social services organization and consulted with a firm specializing in the nonprofit sector. Additionally, we delved into the nonprofit and economics literature to gain insights from both the charities' and volunteers' perspectives. Our proposed model considers two unique features in volunteer management. First, contrary to the literature that assumes volunteers are homogeneous, we characterize volunteers based on their turnout reliability and work performance. Second, we consider volunteers' value in both labor contribution and monetary donations, where the endogenous decision of team composition partially influences monetary donation.

We study two variants of volunteer scheduling models; First, we present a labor-only objective model (VS-L) that provides a baseline, which resembles charity's current practice. Next, we present VS-J model that considers volunteers' labor and donations, but the model is more complex as it is neither concave nor concave. We obtain closed-form expressions for both models by assuming a uniform distribution of episodic volunteers' turnout probability. Results show that although formal volunteers are always preferred and prioritized in the VS-L model, episodic volunteers can serve as the primary workforce when charity considers their monetary donations, too. Moreover, we obtain additional insights regarding how the reliability of episodic volunteers influence the optimal policy and generalize implications for process improvements.

We also applied our model to various volunteering tasks within a social services organization. Given the challenges related to data quality, we devised an additional method to derive an optimal policy while accommodating distribution ambiguity. Our numerical experiments demonstrate that charities can potentially reduce their profit loss by an average of 8.87% with our proposed DRO policy compared to their existing scheduling process. It is important to note that, in some cases, there is still significant room for improvement, even with the DRO approach. This underscores the crucial role of a reliable data infrastructure, which can enhance both volunteer and donor management for charities. Lastly, we created an intuitive Excel-based volunteer staffing tool and employed an interpretable machine learning model to generate actionable insights in simple terms.

Finally, this study also paves the way for future research opportunities. For instance, while our solution approach is tailored to address the data challenges commonly encountered by charities, we firmly believe that a finely-tuned parametric model could outperform our solution as charities establish more rigorous data collection processes. Secondly, enhancing the data collection process itself presents an interesting and rewarding challenge. It is crucial for charities to maintain positive relationships with their volunteers during data collection. Consequently, future researchers can explore methods to track attendance data in a friendly and non-intrusive manner.

References

Ai W, Chen R, Chen Y, Mei Q, Phillips W (2016) Recommending teams promotes prosocial lending in online microfinance. *Proceedings of the National Academy of Sciences* 113(52):14944–14948.

AmeriCorps (2023) Volunteering and civic life in america. Accessed September 10, 2023, https://americorps.gov/sites/default/files/document/volunteering-civic-life-america-research-summary.pdf.

Andreoni J (1990) Impure altruism and donations to public goods: A theory of warm-glow giving. *The economic journal* 100(401):464–477.

Apinunmahakul A, Barham V, Devlin R (2009) Charitable giving, volunteering, and the paid labor market. *Nonprofit and Voluntary Sector Quarterly* 38(1):77–94.

Ata B, Lee D, Sönmez E (2019) Dynamic volunteer staffing in multicrop gleaning operations. *Operations Research* 67(2):295–314.

Authors (2023) Does volunteering crowd out donations? Evidence from online experiments. *Working Paper* .

Beder J, Fast J (2008) Episodic volunteering: Why people walk/run for charity. *The International Journal of Volunteer Administration* 25(3):3–13.

Bekkers R, Wiepking P (2011) A literature review of empirical studies of philanthropy: Eight mechanisms that drive charitable giving. *Nonprofit and Voluntary Sector Quarterly* 40(5):924–973.

Berman O, Larson RC, Pinker E (1997) Scheduling workforce and workflow in a high volume factory. *Management Science* 43(2):158–172.

Brown AL, Meer J, Williams JF (2019) Why do people volunteer? An experimental analysis of preferences for time donations. *Management Science* 65(4):1455–1468.

Brown E, Lankford H (1992) Gifts of money and gifts of time estimating the effects of tax prices and available time. *Journal of Public Economics* 47(3):321–341.

Brudney J, Meijs L (2009) It ain't natural: Toward a new (natural) resource conceptualization for volunteer management. *Nonprofit and Voluntary Sector Quarterly* 38(4):564–581.

Bugen L (1977) Composition and orientation effects on group cohesion. *Psychological Reports* 40(1):175–181.

Burlando R, Guala F (2005) Heterogeneous agents in public goods experiments. *Experimental Economics* 8(1):35–54.

Cappellari L, Ghinetti P, Turati G (2011) On time and money donations. *The Journal of Socio-Economics* 40(6):853 – 867.

Charness G, Chen Y (2020) Social identity, group behavior, and teams. *Annual Review of Economics* 12:691–713.

Clary EG, Snyder M, Ridge RD, Copeland J, Stukas AA, Haugen J, Miene P (1998) Understanding and assessing the motivations of volunteers: A functional approach. *Journal of Personality and Social Psychology* 74(6):1516.

Cnaan RA, Handy F (2005) Towards understanding episodic volunteering. *Vrijwillige Inzet Onderzocht* 2(1):29–35.

Cnaan RA, Meijs L, Brudney JL, Hersberger-Langloh S (2022) You thought that this would be easy? Seeking an understanding of episodic volunteering. *Voluntas* 33:415–427.

Daniels K, Valdés L (2021) Trying and failing: Biases in donor aversion to rejection. *Production and Operations Management* 30(12):4356 – 4373.

De La Torre Pacheco S, Eftekhar M, Wu C (2023) Improving the quality of in-kind donations: A field experiment. *Manufacturing & Service Operations Management* 25(5):1677–1691.

Dietz R, Keller B (2016) Donor loyalty study: A deep dive into donor behaviors and attitudes. *Nonprofit Times* .

Dwiggins-Beeler R, Spitzberg B, Roesch S (2011) Vectors of volunteerism: Correlates of volunteer retention, recruitment, and job satisfaction. *Journal of Psychological Issues in Organizational Culture* 2(3):22–43.

Ellis J (2007) Best practices in volunteer management: An action planning guide for small and rural nonprofit organizations. Accessed May 30, 2021, https://www.volunteeryukon.ca/uploads/general/Best_Practices_Volunteer_Management.pdf.

Falasca M, Zobel C (2012) An optimization model for volunteer assignments in humanitarian organizations. *Socio-Economic Planning Sciences* 46(4):250 – 260.

Feldman N (2010) Time is money: Choosing between charitable activities. *American Economic Journal: Economic Policy* 2(1):103–30.

Fidelity Charitable (2014) Time and money: The role of volunteering in philanthropy. Accessed June 2, 2021, https://www.fidelitycharitable.org/content/dam/fc-public/docs/insights/volunteering-and-philanthropy.pdf.

Fisher M, Krishnan J, Netessine S (2006) Retail store execution: An empirical study. *Preprint, Available at SSRN: https://dx.doi.org/10.2139/ssrn.2319839* .

Fraser J, Clayton S, Sickler J, Taylor A (2009) Belonging at the zoo: Retired volunteers, conservation activism and collective identity. *Ageing and Society* 29(3):351–368.

Fritz J (2019) Why charities should treat volunteers like donors. *The Balance Small Business* .

Gachter S, Thoni C (2005) Social learning and voluntary cooperation among like-minded people. *Journal of the European Economic Association* 3(2-3):303–314.

Gage R, Thapa B (2012) Volunteer motivations and constraints among college students: Analysis of the volunteer function inventory and leisure constraints models. *Nonprofit and Voluntary Sector Quarterly* 41(3):405–430.

Gordon L, Erkut E (2004) Improving volunteer scheduling for the edmonton folk festival. *INFORMS Journal on Applied Analytics* 34(5):367–376.

Gray D, Stevenson C (2020) How can we help? Exploring the role of shared social identity in the experiences and benefits of volunteering. *Journal of Community & Applied Social Psychology* 30(4):341–353.

Hustinx L, Haski-Leventhal D, Handy F (2008) One of a kind? Comparing episodic and regular volunteers at the philadelphia ronald mcdonald house. *International Journal of Volunteer Administration* 25(3):50–66.

Hustinx L, Lammertyn F (2003) Collective and reflexive styles of volunteering: A sociological modernization perspective. *Voluntas: International Journal of Voluntary and Nonprofit Organizations* 14(2):167–187.

Hyde M, Dunn J, Bax C, Chambers S (2016) Episodic volunteering and retention: An integrated theoretical approach. *Nonprofit and Voluntary Sector Quarterly* 45(1):45–63.

Jami A, Kouchaki M, Gino F (2021) I own, so i help out: How psychological ownership increases prosocial behavior. *Journal of Consumer Research* 47(5):698–715.

Kamalahmadi M, Yu Q, Zhou Y (2021) Call to duty: Just-in-time scheduling in a restaurant chain. *Management Science* 67(11):6751–6781.

Kesavan S, Lambert S, Williams J, Pendem P (2022) Doing well by doing good: Improving retail store performance with responsible scheduling practices at the Gap, Inc. *Management Science* 0(0).

Kesavan S, Staats BR, Gilland W (2014) Volume flexibility in services: The costs and benefits of flexible labor resources. *Management Science* 60(8):1884–1906.

Lassiter K, Khademi A, Taaffe KM (2015) A robust optimization approach to volunteer management in humanitarian crises. *International Journal of Production Economics* 163:97 – 111.

Lieberman M, Wizlenberg A, Golant M, Di Minno M (2005) The impact of group composition on internet support groups: Homogeneous versus heterogeneous parkinson's groups. *Group Dynamics: Theory, Research, and Practice* 9(4):239.

Liu W, Aaker J (2008) The happiness of giving: The time-ask effect. *Journal of Consumer Research* 35(3):543–557.

Low N, Butt S, Ellis Paine A, Davis Smith J (2007) Helping out: A national survey of volunteering and charitable giving.

Manshadi V, Rodilitz S (2022) Online policies for efficient volunteer crowdsourcing. *Management Science* 0(0).

Mason AJ, Ryan DM, Panton DM (1998) Integrated simulation, heuristic and optimization approaches to staff scheduling. *Operations Research* 46(2):161–175.

Miller LE, Powell GN, Seltzer J (1990) Determinants of turnover among volunteers. *Human Relations* 43(9):901–917.

Olsen JA, Eidem JI (2003) An inquiry into the size of health charities: The case of Norwegian patient organizations. *Journal of Socio-Economics* 32:457–466.

Parsa I, Eftekhar M, Corbett JC (2022) Does better governance relax the efficiency squeeze experienced by non-profit organizations? *Production and Operations Management* 31(8):3288–3303.

Peck J, Kirk C, Luangrath A, Shu S (2021) Caring for the commons: using psychological ownership to enhance stewardship behavior for public goods. *Journal of Marketing* 85(2):33–49.

Pierce J, Jussila I (2010) Collective psychological ownership within the work and organizational context: Construct introduction and elaboration. *Journal of Organizational Behavior* 31(6):810–834.

Polik I, Terlaky T (2007) A survey of the s-lemma. *SIAM Review* 49(3):371–418.

Productivity Commission (2008) Contribution of the not-for-profit sector. Technical report, Canberra: Australian Government.

Sampson S (2006) Optimization of volunteer labor assignments. *Journal of Operations Management* 24(4):363–377.

Scarf H (1958) A min-max solution of an inventory problem. *Studies in the mathematical theory of inventory and production* .

Smilowitz K, Nowak M, Jiang T (2013) Workforce management in periodic delivery operations. *Transportation Science* 47(2):214–230.

Smith JD (1998) *The 1997 national survey of volunteering* (National centre for Volunteering).

Ulmer MW, Savelsbergh M (2020) Workforce scheduling in the era of crowdsourced delivery. *Transportation Science* 54(4):1113–1133.

Urrea G, Pedraza-Martinez AJ, Besiou M (2019) Volunteer management in charity storehouses: Experience, congestion and operational performance. *Production and Operations Management* 28(10):2653–2671.

Van den Bergh J, Belien J, De Bruecker P, Demeulemeester E, De Boeck L (2013) Personal scheduling: A literature review. *European Journal of Operational Research* 226(3):367–385.

Vantilborgh T, Bidee J, Pepermans R, Willems J, Huybrechts G, Jegers M (2011) A new deal for npo governance and management: Implications for volunteers using psychological contract theory. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations* 22(4):639–657.

Vantilborgh T, Bidee J, Pepermans R, Willems J, Huybrechts G, Jegers M (2012) Volunteers' psychological contracts: Extending traditional views. *Nonprofit and Voluntary Sector Quarterly* 41(6):1072–1091.

Vecina ML, Chacon F, Sueiro M, A B (2012) Volunteer engagement: Does engagement predict the degree of satisfaction among new volunteers and the commitment of those who have been active longer? *Applied Psychology* 61:130–148.

Villarreal MC, Goldsman D, Keskinocak P (2015) Workforce management and scheduling under flexible demand. *Service Science* 7(4):331–351.

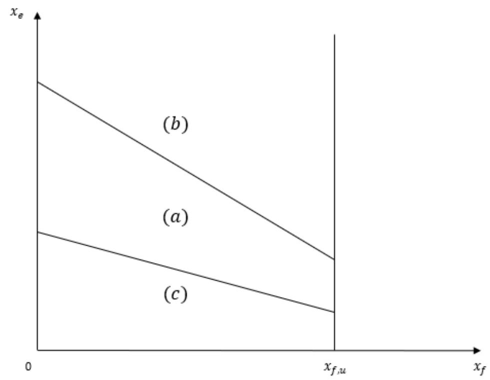


Figure 7 **Subregions of** $L(x)$

Appendix A: Proof of Theorem 1

Proof. Given the composite of volunteer mix (x_e, x_f), equation VS-L has three subregions, as displayed in Figure 7. We analyze each subregion separately.

$$L(\mathbf{x}) = \begin{cases} w\lambda - \gamma(x_e(\frac{h_0+h_u}{2}) + \theta x_f - \lambda)(\frac{h_u-h_0}{h_u-h_l}) - \beta(\lambda - \theta x_f - (\frac{h_0+h_l}{2})x_e)(\frac{h_0-h_l}{h_u-h_l}), & \text{case (a): if } h_l < h_0 < h_u \\ w\lambda - \gamma(x_e(\frac{h_l+h_u}{2}) + \theta x_f - \lambda), & \text{case (b): if } h_0 \le h_l \\ w\lambda - \beta(\lambda - \theta x_f - (\frac{h_u+h_l}{2})x_e), & \text{case (c): if } h_0 \ge h_u \end{cases}$$

Case (a) In this region, we can check that the Hessian of L is negative semi-definite, and hence $L(\mathbf{x})$ is jointly concave in $\mathbf{x}$. Analyzing the gradient

$$\nabla L(\mathbf{x}) = \left(\frac{(\beta+\gamma)(\lambda-\theta x_f)^2 - (h_l^2\beta + h_u^2\gamma)x_e^2}{2x_e^2(h_u-h_l)}, \ \frac{\theta(\beta+\gamma)(\lambda-\theta x_f - (h_l+h_u)x_e}{(h_u-h_l)x_e} \right),$$

we observe that $\nabla L(\mathbf{x}) = 0$ does not have a solution. Hence, the solution to $\max_{\mathbf{x}} L(\mathbf{x})$, which we denote as $\mathbf{x}^* = (x_e^*, x_f^*)$, must lie on the boundary of region X. We will next derive the optimal value of x_e (x_e^*) for a given x_f. The optimal x_e^* can be derived by solving the first order condition, $\frac{(\beta+\gamma)(\lambda-\theta x_f)^2-(h_l^2\beta+h_u^2\gamma)x_e^2}{2x_e^2(h_u-h_l)} = 0$. The optimal $x_e = \frac{(\lambda-\theta x_f^*)\sqrt{\beta+\gamma}}{\sqrt{h_l^2\beta+h_u^2\gamma}}$. Therefore, the objective function value is as follows.

$$L(x_e^*, x_f) = w\lambda + \frac{(\lambda-\theta x_f)(h_l\beta + h_u\gamma - \sqrt{(\beta+\gamma)(h_l^2\beta+h_u^2\gamma)}}{h_u-h_l}$$

Note that $h_l\beta + h_u\gamma - \sqrt{(\beta+\gamma)(h_l^2\beta+h_u^2\gamma)} < 0$ since $h_l\beta + h_u\gamma = \sqrt{h_l^2\beta^2 + h_u^2\gamma^2 + 2h_lh_u\gamma\beta} < \sqrt{h_l^2\beta^2 + h_u^2\gamma^2 + (h_l^2+h_u^2)\beta\gamma} = \sqrt{(\beta+\gamma)(h_l^2\beta+h_u^2\gamma)}$. Therefore, $L(x_e^*, x_f)$ increases in x_f. So the optimal $x_f^* = \bar{x}_f = \min\{x_{f,u}, \frac{\lambda}{\theta}\}$, and the optimal objective function is $L^* = w\lambda + \frac{(\lambda-\theta x_f)(\beta h_l + h_u\gamma - \sqrt{\beta+\gamma}\sqrt{h_l^2\beta+h_u^2\gamma})}{h_u-h_l}$.

Cases (b) and (c) These two cases will not yield global optimal solutions. Note that the expected turnout of episodic volunteers h is $\frac{h_l+h_u}{2}$. In case (b) (c), the policy is to invite an insufficient (excessive) number of volunteers and the charity is expected to have a shortage of volunteers (surplus). In either case, the charity can improve its performance by increasing or decreasing the number of episodic volunteers it wants to invite. Therefore, the local optimal solution of cases (b) and (c) can exist only on the line $h_0 = h_l$ or $h_0 = h_u$. Furthermore, $L(\mathbf{x}) = w\lambda - \gamma(x_e\frac{h_l+h_u}{2} + \theta x_f - \lambda)$ when $h_0 = h_l$ and $L(\mathbf{x}) = w\lambda - \beta(\lambda - x_e\frac{h_l+h_u}{2} - \theta x_f)$ when $h_0 = h_u$. In both cases, $L(\mathbf{x})$ is linear in x_f, so the optimal solution must be on the lines $x_f = 0$ or $x_f = \bar{x}_f$.

■

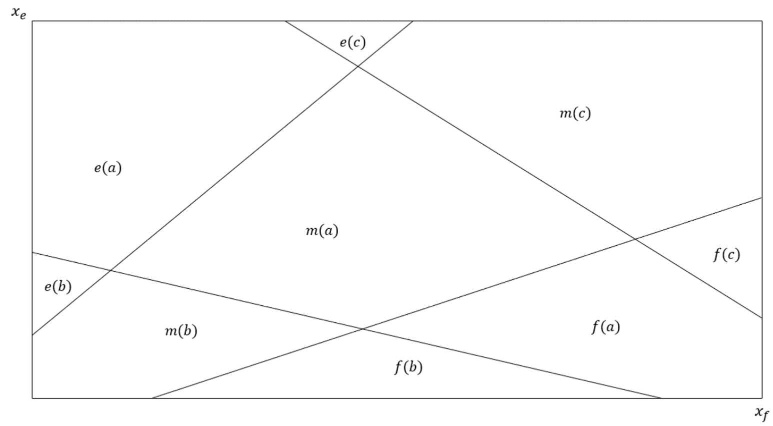


Figure 8 **Subregions of $\mathcal{X}$**

Appendix B: Proof of Theorem 2 and 3

Proof. Given the closed-form expression for $J(\mathbf{x})$, we will solve equation VS-J that maximizes $J(\mathbf{x})$ in the feasible region $\mathcal{X}$. Since $x_{f,\ell}=0$ and $x_{f,u}\geq\lambda/\theta$, the constraint $x_f\in[x_{f,\ell},x_{f,u}]$ in $\mathcal{X}$ is redundant. In total, we have nine cases for the closed-form of $J(\mathbf{x})$. We define the following regions of $\mathcal{X}$:

$$X_e := \left\{\mathbf{x}=(x_e,x_f)\in\mathbb{R}\times\mathbb{R}\ :\ x_f,x_e\geq 0,\ h_f:=\frac{\alpha x_f}{x_e}<h_l\right\}$$
$$X_m := \left\{\mathbf{x}=(x_e,x_f)\in\mathbb{R}\times\mathbb{R}\ :\ x_f,x_e\geq 0,\ h_f:=\frac{\alpha x_f}{x_e}\in[h_l,h_u]\right\}$$
$$X_f := \left\{\mathbf{x}=(x_e,x_f)\in\mathbb{R}\times\mathbb{R}\ :\ x_f,x_e\geq 0,\ h_f:=\frac{\alpha x_f}{x_e}>h_u\right\}$$

These regions are illustrated in Figure 8 and labeled as 'e', 'm', and 'f'. Note that all three regions X_e, X_m, and X_f are polyhedra, as they are defined by linear constraints. If $\mathbf{x}\in X_f$ ($\mathbf{x}\in X_m$, $\mathbf{x}\in X_e$), then the closed-form expression of $J(\mathbf{x})$ corresponds to cases f(a)-f(c) (cases m(a)-m(c), e(a)-e(c)).

Given a subregion label ℓ, we will use J_ℓ to refer to its closed-form expression. For example, if $\mathbf{x}\in X_{m(a)}$, then $J(\mathbf{x})=J_{m(a)}(\mathbf{x})$ where $J_{m(a)}(\mathbf{x}):=w\lambda-\gamma(x_e(\frac{h_0+h_u}{2})+\theta x_f-\lambda)(\frac{h_u-h_0}{h_u-h_l})-\beta(\lambda-\theta x_f-(\frac{h_0+h_l}{2})x_e)(\frac{h_0-h_l}{h_u-h_l})+d_e(\frac{h_u+h_l}{2})x_e+d_f x_f+d'_f x_f(\alpha-\frac{x_e}{x_f}(\frac{h_f+h_l}{2}))\frac{h_f-h_l}{h_u-h_l}$. We similarly define functions $J_{e(i)}$, $J_{m(i)}$ and $J_{f(i)}$ for $i\in\{a,b,c\}$.

Our goal is to find the global maximizer of $J(\mathbf{x})$ among all subregions of $\mathcal{X}$. Hence, to solve the maximization problem in equation VS-J, we should analyze the maximum of $J_e(\cdot)$, $J_m(\cdot)$, and $J_f(\cdot)$, in all subregions of X_e, X_m, and X_f, respectively. We solve the problem in two steps. In the first step, we will show that maximizing $J(\mathbf{x})$ over each subregion achieves a solution that is at the boundary of the subregion. In the second step, we will compare the values of J on the boundary of each subregion to derive the optimal global solution.

Step 1: We will prove that the local maximizer of J within each subregion of X_e, X_m, and X_f is at the boundary of the subregion. These boundaries are illustrated in Figure 8. We use the notation $\mathrm{bd}(X)$ to refer to the boundary of a region X.

Subregion $X_{e(a)}$. When $h_f<h_l$, the objective function becomes $J_{e(a)}=w\lambda-\gamma(hx_e+\theta x_f-\lambda)^+-\beta(\lambda-hx_e-\theta x_f)^++d_e hx_e+d_f x_f=L(\mathbf{x})+d_e hx_e+d_f x_f$. Since $J_{e(a)}$ only has two more linear terms in addition

to L, we must have $\nabla J_{e(a)} = \nabla L = 0$, and the optimal solution to $\max_{\mathbf{x}\in X_{e(a)}} J(\mathbf{x})$ must lie on the boundary $\text{bd}\left(X_{e(a)}\right)$.

Subregion $X_{m(a)}$. When $h_f \in [h_l, h_u]$, the objective becomes:

$$\begin{aligned}\mathbb{E}[J_{m(a)}] &= \mathbb{E}\left[w\lambda - \gamma(hx_e + \theta x_f - \lambda)^+ - \beta(\lambda - hx_e - \theta x_f)^+ + d_e hx_e + d_f x_f + d'_f x_f(\alpha - \frac{hx_e}{x_f})^+\right] \\ &= w\lambda - \gamma(x_e(\frac{h_0+h_u}{2}) + \theta x_f - \lambda)(\frac{h_u-h_0}{h_u-h_l}) - \beta(\lambda - \theta x_f - (\frac{h_0+h_l}{2})x_e)(\frac{h_0-h_l}{h_u-h_l}) \\ &+ d_e(\frac{h_u+h_l}{2})x_e + d_f x_f + d'_f x_f(\alpha - \frac{x_e}{x_f}(\frac{h_f+h_l}{2}))\frac{h_f-h_l}{h_u-h_l}\end{aligned}$$

Next, the second-order derivative and cross partial derivative are as follow: $\frac{\partial^2 E(J_{m(a)})}{\partial x_e^2} = \frac{d'_f x_f^2\alpha^2 - (\beta+\gamma)(\lambda-\theta x_f)^2}{x_e^3(h_u-h_l)}$, $\frac{\partial^2 E(J_{m(a)})}{\partial x_f^2} = \frac{d'_f\alpha^2 - (\beta+\gamma)\theta^2}{x_e(h_u-h_l)}$, and $\frac{\partial^2 E(J_{m(a)})}{\partial x_f\partial x_e} = \frac{-d'_f x_f\alpha^2 - (\beta+\gamma)\theta(\lambda-\theta x_f)}{x_e^2(h_u-h_l)}$. Moreover, $\det(H(E(J_{m(a)}))) = -\frac{d'_f\alpha^2(\beta+\gamma)\lambda^2}{x_e^4(h_u-h_l)^2} < 0$

Since the determinant of the Hessian is always negative, the solution of $\max_{\mathbf{x}\in X_{m(a)}} J(\mathbf{x})$ is at the boundary $\text{bd}\left(X_{m(a)}\right)$.

Subregion $X_{f(a)}$. When $h_f > h_u$, the objective function becomes $J_{f(a)} = w\lambda - \gamma(hx_e + \theta x_f - \lambda)^+ - \beta(\lambda - hx_e - \theta x_f)^+ + d_e hx_e + d_f x_f + d'_f x_f(\alpha - \frac{hx_e}{x_f}) = L(\mathbf{x}) + d_e hx_e + d_f x_f + d'_f x_f(\alpha - \frac{hx_e}{x_f})$. Since $J_{f(a)}$ only has three more linear terms in addition to L, we must have $\nabla J_{f(a)} = \nabla L = 0$, and the optimal solution to $\max_{\mathbf{x}\in X_{f(a)}} J(\mathbf{x})$ must lie on the boundary $\text{bd}\left(X_{f(a)}\right)$.

Subregion $X_{e(b)}$. When $h_f < h_l$, the objective function becomes $J_{e(b)} = w\lambda - \beta(\lambda - hx_e - \theta x_f) + d_e hx_e + d_f x_f$. Since $J_{e(b)}$ has a linear objective function, the optimal solution to $\max_{\mathbf{x}\in X_{e(b)}} J(\mathbf{x})$ is always at the extreme points.

Subregion $X_{m(b)}$. When $h_l < h_f < h_u$, the objective function becomes:

$$\begin{aligned}E[J_{m(b)}] &= w\lambda - \beta(\lambda - hx_e - \theta x_f) + d_e hx_e + d_f x_f + d'_f x_f(\alpha - \frac{hx_e}{x_f}) \\ &= w\lambda - \beta(\lambda - x_e(\frac{h_0+h_u}{2}) - \theta x_f) + d_e(\frac{h_u+h_l}{2})x_e + d_f x_f + d'_f x_f(\alpha - \frac{x_e}{x_f}(\frac{h_f+h_l}{2}))\frac{h_f-h_l}{h_u-h_l}\end{aligned}$$

Next, the second-order derivative and cross partial derivative are as follows: $\frac{\partial^2 E(J_{m(b)})}{\partial x_e^2} = \frac{d'_f x_f^2\alpha^2}{x_e^3(h_u-h_l)}$, $\frac{\partial^2 E(J_{m(b)})}{\partial x_f^2} = \frac{d'_f\alpha^2}{x_e(h_u-h_l)}$, and $\frac{\partial^2 E(J_{m(b)})}{\partial x_f\partial x_e} = \frac{-d'_f x_f\alpha^2}{x_e^2(h_u-h_l)}$. Since $\nabla E(J_{m(b)}) = 0$, the optimal solution to $\max_{\mathbf{x}\in X_{m(b)}} J(\mathbf{x})$ must be in the boundary $\text{bd}\left(X_{m(b)}\right)$.

Subregion $X_{f(b)}$. When $h_f > h_u$, the objective function becomes $J_{f(b)} = w\lambda - \beta(\lambda - hx_e - \theta x_f) + d_e hx_e + d_f x_f + d'_f x_f(\alpha - \frac{hx_e}{x_f})$. Since $J_{f(b)}$ is linear, the optimal solution to $\max_{\mathbf{x}\in X_{f(b)}} J(\mathbf{x})$ is always at the extreme points.

Subregion $X_{e(c)}$. When $h_f < h_l$, the objective function becomes $J_{e(c)} = w\lambda - \gamma(hx_e + \theta x_f - \lambda) + d_e hx_e + d_f x_f$. Since $J_{e(c)}$ has a linear objective function, the optimal solution to $\max_{\mathbf{x}\in X_{e(c)}} J(\mathbf{x})$ is always at the extreme points.

Subregion $X_{m(c)}$. When $h_l < h_f < h_u$, the objective function becomes:

$$\begin{aligned}E[J_{m(c)}] &= w\lambda - \gamma(hx_e + \theta x_f - \lambda) + d_e hx_e + d_f x_f + d'_f x_f(\alpha - \frac{hx_e}{x_f}) \\ &= w\lambda - \gamma(x_e(\frac{h_l+h_u}{2}) + \theta x_f - \lambda) + d_e(\frac{h_u+h_l}{2})x_e + d_f x_f + d'_f x_f(\alpha - \frac{x_e}{x_f}(\frac{h_f+h_l}{2}))\frac{h_f-h_l}{h_u-h_l}\end{aligned}$$

Since $\nabla E(J_{m(c)}) = 0$, the optimal solution to $\max_{\mathbf{x}\in X_{m(b)}} J(\mathbf{x})$ must lie on the boundary $\text{bd}\left(X_{m(c)}\right)$.

Subregion $X_{f(c)}$. When $h_f > h_u$, the objective function becomes $J_{f(c)} = w\lambda - \gamma(hx_e + \theta x_f - \lambda) + d_e hx_e + d_f x_f + d'_f x_f(\alpha - \frac{hx_e}{x_f})$. Since $J_{f(c)}$ is linear, the optimal solution to $\max_{\mathbf{x}\in X_{f(c)}} J(\mathbf{x})$ is always at the extreme points.

Step 2: Next, we analyze the value of $J(\mathbf{x})$ on the candidate optimal solutions identified in Step 1. Note that these candidate solutions are the boundaries of the subregions in X_e, X_m, and X_f in Figure 8. In this step, we further eliminate boundaries as candidate solutions.

Boundary lines $\text{bd}(X_{m(a)}) \cap \text{bd}(X_{e(a)})$ and $\text{bd}(X_{m(a)}) \cap \text{bd}(X_{f(a)})$.

$$\frac{\partial E(J_{m(a)})}{\partial x_f} = d_f + \frac{d'_f\alpha(x_f\alpha - h_l x_e) + \theta((\beta+\gamma)(\lambda - x_f\theta) - h_l x_e\beta - h_u x_e\gamma)}{x_e(h_u - h_l)}$$
$$\frac{\partial E(J_{e(a)})}{\partial x_f} = d_f + \frac{\theta((\beta+\gamma)(\lambda - x_f\theta) - h_l x_e\beta - h_u x_e\gamma)}{x_e(h_u - h_l)}$$
$$\frac{\partial E(J_{f(a)})}{\partial x_f} = d_f + d'_f\alpha + \frac{\theta((\beta+\gamma)(\lambda - x_f\theta) - h_l x_e\beta - h_u x_e\gamma)}{x_e(h_u - h_l)}$$

If the optimal solution lies on the boundary line $\text{bd}(X_{m(a)}) \cap \text{bd}(X_{e(a)})$, then we must have $\frac{\partial E(J_{e(a)})}{\partial x_f} > 0$ and $\frac{\partial E(J_{m(a)})}{\partial x_f} < 0$. However, $\frac{\partial E(J_{m(a)})}{\partial x_f} > \frac{\partial E(J_{e(a)})}{\partial x_f}$ in $X_{m(a)}$.

Similarly, if the optimal solution lies on the boundary line $\text{bd}(X_{m(a)}) \cap \text{bd}(X_{f(a)})$, then we must have $\frac{\partial E(J_{f(a)})}{\partial x_f} < 0$ and $\frac{\partial E(J_{m(a)})}{\partial x_f} > 0$. However, $\frac{\partial E(J_{m(a)})}{\partial x_f} < \frac{\partial E(J_{f(a)})}{\partial x_f}$ in $X_{m(a)}$. Therefore, the optimal solution cannot exist on these two boundary lines.

Boundary lines $\text{bd}(X_{m(b)}) \cap \text{bd}(X_{e(b)})$ and $\text{bd}(X_{m(b)}) \cap \text{bd}(X_{f(b)})$.

$$\frac{\partial E(J_{m(b)})}{\partial x_f} = d_f + \beta\theta + \frac{d'_f\alpha(x_f\alpha - h_l x_e)}{x_e(h_u - h_l)}$$
$$\frac{\partial E(J_{e(b)})}{\partial x_f} = d_f + \beta\theta$$
$$\frac{\partial E(J_{f(b)})}{\partial x_f} = d_f + d'_f\alpha + \beta\theta$$

If the optimal solution lies on the boundary line $\text{bd}(X_{m(b)}) \cap \text{bd}(X_{e(b)})$, then we must have $\frac{\partial E(J_{e(b)})}{\partial x_f} > 0$ and $\frac{\partial E(J_{m(b)})}{\partial x_f} < 0$. However, $\frac{\partial E(J_{m(b)})}{\partial x_f} > \frac{\partial E(J_{e(b)})}{\partial x_f}$ in $X_{m(b)}$.

Similarly, if the optimal solution lies on the boundary line $\text{bd}(X_{m(b)}) \cap \text{bd}(X_{f(b)})$, then we must have $\frac{\partial E(J_{f(b)})}{\partial x_f} < 0$ and $\frac{\partial E(J_{m(b)})}{\partial x_f} > 0$. However, $\frac{\partial E(J_{f(b)})}{\partial x_f} > \frac{\partial E(J_{m(b)})}{\partial x_f}$ in $X_{m(b)}$. Therefore, the optimal solution cannot exist on these two boundary lines.

Boundary lines $\text{bd}(X_{m(c)}) \cap \text{bd}(X_{e(c)})$ and $\text{bd}(X_{m(c)}) \cap \text{bd}(X_{f(c)})$.

$$\frac{\partial E(J_{m(c)})}{\partial x_f} = d_f - \gamma\theta + \frac{d'_f\alpha(x_f\alpha - h_l x_e)}{x_e(h_u - h_l)}$$
$$\frac{\partial E(J_{e(c)})}{\partial x_f} = d_f - \gamma\theta$$
$$\frac{\partial E(J_{f(c)})}{\partial x_f} = d_f + d'_f\alpha - \gamma\theta$$

If the optimal solution lies on the boundary line $\text{bd}(X_{m(c)}) \cap \text{bd}(X_{e(c)})$, then we must have $\frac{\partial E(J_{e(c)})}{\partial x_f} > 0$ and $\frac{\partial E(J_{m(c)})}{\partial x_f} < 0$. However, $\frac{\partial E(J_{m(c)})}{\partial x_f} > \frac{\partial E(J_{e(c)})}{\partial x_f}$ in $X_{m(c)}$.

Similarly, if the optimal solution lies on the boundary line $\text{bd}(X_{m(c)}) \cap \text{bd}(X_{f(c)})$, then we must have $\frac{\partial E(J_{f(c)})}{\partial x_f} < 0$ and $\frac{\partial E(J_{m(c)})}{\partial x_f} > 0$. However, $\frac{\partial E(J_{f(c)})}{\partial x_f} > \frac{\partial E(J_{m(c)})}{\partial x_f}$ in $X_{m(c)}$. Therefore, an optimal solution cannot exist on these two boundary lines.

<u>Boundary lines $h_0 = h_l$ $(\text{bd}(X_{m(c)}) \cap \text{bd}(X_{m(a)}))$ and $h_0 = h_u$ $(\text{bd}(X_{m(b)}) \cap \text{bd}(X_{f(a)}))$.</u>

Since $J_{e(b)}, J_{f(b)}, J_{e(c)}, J_{f(c)}$ all have linear objective functions, the maxima of these regions must reside at the extreme points. Furthermore, the objective functions of the lines $X_m(c) \cap X_m(a)$ and $X_m(b) \cap X_m(a)$ are convex. On the line $h_0 = h_l$, the second-order derivative of $\frac{\partial^2 J_{m(a)}}{\partial x_f^2} = \frac{d'_f h_l \alpha^2 \lambda^2}{(h_u - h_l)(\lambda - \theta x_f)^3} > 0$. On the line $h_0 = h_u$, the second order derivative of $\frac{\partial^2 J_{m(a)}}{\partial x_f^2} = \frac{d'_f h_u \alpha^2 \lambda^2}{(h_u - h_l)(\lambda - \theta x_f)^3} > 0$. Therefore, both objective functions are convex in x_f and the maxima must be on the boundary. Last, since we already proved that the optimal solution will not be on the line $h_f = h_l$ or the line $h_f = h_u$, the maxima can only exist on $x_e = 0$, $x_f = 0$.

<u>Boundary lines $x_e = 0$ and $x_f = 0$.</u>

First, when $x_e = 0$, the objective function becomes deterministic, and the optimal solution and objective functions are $x_f^* = \lambda/\theta$ and $J(0, \lambda/\theta)] = w\lambda + \frac{(d_f + \alpha d'_f)\lambda}{\theta}$ when $d_f + d'_f \alpha < \theta\gamma$, and $x_f^* = x_{f,u}$, $J(0, x_{f,u})] = (d_f + \alpha d'_f) x_{f,u} - \gamma(\theta x_{f,u} - \lambda)$ otherwise.

Second, when $x_f = 0$, the objective function becomes: $\mathbb{E}[J(x_e, 0)] = w\lambda - \gamma(hx_e - \lambda)^+ - \beta(\lambda - hx_e)^+ + d_e h x_e$. The optimal solution is $x_e^* = x_{e,1}^* = \frac{\lambda\sqrt{\beta+\gamma}}{\sqrt{d_e(h_l^2 - h_u^2) + h_l^2\beta + h_u^2\gamma}}$ when $d_e \le \gamma$, and $x_e^* = x_{e,u}$ otherwise. Note that the condition for $x_{e,1}^*$ to hold is $d_e < \gamma$ $d_e < \frac{h_u^2\gamma + h_l^2\beta}{h_u^2 - h_l^2}$ which could be greater than γ. The objective function $J_e(m)$ is always dominated by $J_e(c)$ when $d_e > \gamma$. Therefore, $x_{e,1}^*$ cannot be the optimal solution when $d_e > \gamma$. The objective function value of $(x_{e,1}^*, 0)$ is $\mathbb{E}[J(x_{e,1}^*, 0)] = w\lambda + \frac{\lambda(h_l\beta + h_u\gamma - \sqrt{\beta+\gamma}\sqrt{d_e(h_l^2 - h_u^2) + h_l^2\beta + h_u^2\gamma})}{h_u - h_l} = w\lambda + \lambda\gamma + \frac{h_l(\beta-\gamma)}{h_u - h_l} - \frac{\sqrt{\beta+\gamma}\sqrt{h_l^2\beta + h_u^2\gamma - (h_u^2 - h_l^2)d_e}}{h_u - h_l}$. The objective function of $(x_{e,u}, 0)$ is $\mathbb{E}[J(x_{e,u}, 0)] = w\lambda - \gamma(\frac{h_l + h_u}{2} x_{e,u} - \lambda) + d_e \frac{h_l + h_u}{2} x_{e,u}$.

<u>Boundary lines $x_e = x_{e,u}$ and $x_f = x_{f,u}$.</u>

On the line $x_e = x_{e,u}$ and the line $x_f = x_{f,u}$, both $J_{e(c)}$ and $J_{f(c)}$ are linear in (x_e, x_f). Furthermore, $J_{m(c)}$ is convex in both x_e and x_f. Therefore, the optimal solution can only be among the following: $(x_{e,u}, x_{f,u}), (0, x_{f,u}), (x_{e,u}, 0)$. The objective function of $(x_{e,u}, x_{f,u})$ can be either $J_{m(c)}$ or $J_{f(c)}$. If $(x_{e,u}, x_{f,u}) \in X_{m(c)}$, $E[J_{m(c)}] = w\lambda - \gamma(x_{e,u}(\frac{h_l + h_u}{2}) + \theta x_{f,u} - \lambda) + d_e(\frac{h_u + h_l}{2}) x_{e,u} + d_f x_{f,u} + d'_f x_{f,u}(\alpha - \frac{x_{e,u}}{x_{f,u}}(\frac{h_f + h_l}{2})) \frac{h_f - h_l}{h_u - h_l}$ where $h_0 = \frac{\lambda - \theta x_{f,u}}{x_{e,u}}$ and $h_f = \frac{\alpha x_{f,u}}{x_{e,u}}$. If $(x_{e,u}, x_{f,u}) \in X_{f(c)}$, $E[J_{f(c)}] = w\lambda - \gamma(\frac{h_l + h_u}{2} x_{e,u} + \theta x_{f,u} - \lambda) + d_e \frac{h_l + h_u}{2} x_{e,u} + d_f x_{f,u} + d'_f x_{f,u}(\alpha - \frac{\frac{h_l + h_u}{2} x_{e,u}}{x_{f,u}})$.

<u>Conclusion:</u>

Last, we compare the optimal conditions and objective values at each point to summarize the final results in Table 3.

- When $d_f < \theta\gamma - \alpha d'_f$ and $d_e < \gamma$, there are only two possible optimal solutions: $(x_{e,1}^*, 0)$, $(0, \lambda/\theta)$. The optimal solution is $(x_{e,1}^*, 0)$ if $\nu = \frac{(h_l\beta + h_u\gamma - \sqrt{\beta+\gamma}\sqrt{d_e(h_l^2 - h_u^2) + h_l^2\beta + h_u^2\gamma})}{h_u - h_l} > (d_f + \alpha d'_f)/\theta$, and the optimal solution is $(0, \lambda/\theta)$ otherwise. This completes the proof for Theorem 3.

- when $d_f < \theta\gamma - \alpha d'_f$ and $d_e > \gamma$, the optimal solution is $(x_{e,u}, 0)$ if $J(x_{e,u}, 0) > (d_f + \alpha d'_f)\lambda/\theta$, and the optimal solution is $(0, \lambda/\theta)$ otherwise.

- when $d_f > \theta\gamma - \alpha d'_f$ and $d_e < \gamma$, the only optimal solution is $(0, x_{f,u})$ because $J(0, x_{f,u})^* \ge \max(J(x_{e,1}^*, 0), J(0, \lambda/\theta))$ always holds under this condition. First, $J_2(x_{f,u})^* \ge J_2(\lambda/\theta)$ because $d_f >$

Table 3 Solution for Uniform Distribution

	$d_f < \theta\gamma - \alpha d'_f$	$\theta\gamma - \alpha d'_f \le d_f \le \theta\gamma$	$d_f > \theta\gamma$
$d_e < \gamma$	$(x^*_{e,1}, 0)$ or $(0, \frac{\lambda}{\theta})$	$(0, x_{f,u})$	
$d_e > \gamma$	$(x_{e,u}, 0)$ or $(0, \frac{\lambda}{\theta})$	$(x_{e,u}, 0)$ or $(0, x_{f,u})$	$(x_{e,u}, x_{f,u})$ or $(0, x_{f,u})$

$\theta\gamma - \alpha d'_f$. Second, note that $J_1(x^*_{e,1})$ is monotone increasing in d_e and reaches the maximum value when $d_e = \frac{h_u^2\gamma + h_l^2\beta}{h_u^2 - h_l^2}$. Substituting the value into $J(x^*_{e,1}, 0)$, we have $J(x^*_{e,1}, 0) = w\lambda + \frac{\lambda(h_l\beta + h_u\gamma)}{h_u - h_l} \le w\lambda + \gamma\lambda$. Next, note that $J(0, x_{f,u}) = w\lambda + (d'_f\alpha + d_f - \gamma)x_{f,u} + \lambda\gamma \ge w\lambda + \gamma\lambda \ge J_1(x^*_{e,1})$ when $d'_f\alpha + d_f \ge \gamma$. Therefore, when $d'_f\alpha + d_f \ge \gamma$ and $d_e < \gamma$, $J(0, x_{f,u})^* \ge J(x^*_{e,1}, 0)$ always hold.

- When $d_e > \gamma$ and $d_f > \theta\gamma - \alpha d'_f$, the policy $(x^*_{e,1}, 0)$ is not feasible and the policy $(0, \lambda/\theta)$ always results in a lower value than $(0, x_{f,u})$. When $d_f > \theta\gamma$, there are two possible optimal policies: $(x_{e,u}, x_{f,u})$ or $(0, x_{f,u})$. The optimal solution is $(x_{e,u}, x_{f,u})$ if $J(x_{e,u}, x_{f,u}) > J(0, x_{f,u})$ and the optimal solution is $(0, x_{f,u})$ otherwise. Last, when $d_f \in (\theta\gamma - \alpha d'_f, \theta\gamma)$, there are also two possible optimal policies: $(x_{e,u}, 0)$ and $(0, x_{f,u})$. The optimal solution is $(x_{e,u}, 0)$ if $J(x_{e,u}, 0) > J(0, x_{f,u})$ and the optimal solution is $(0, x_{f,u})$ otherwise.

■

Appendix C: DRO solution - Computational method.

Recall $J^x(y,z) := L^x(y,z) + M^x(y,z)$. To develop our solution approach, we will utilize the fact that, for a given $\mathbf{x}$, the moment problem $J_{DRO}(\mathbf{x}) = \inf_{p\in\mathcal{P}} \mathbb{E}_p[J^x(Hx_e, x_f)]$ is a semi-infinite linear program since $p : [h_\ell, h_u] \mapsto \Re^+$ can be viewed as a vector of infinitely many decision variables. The dual of the moment problem is:

$$\begin{aligned} \max_{t,s,r} \quad & t + \mu s + (\mu^2 + \sigma^2) r \\ \text{s.t.} \quad & t + hs + h^2 r \le J^x(hx_e, x_f), \quad h \in [h_\ell, h_u] \end{aligned} \tag{7}$$

This is also a semi-infinite linear program but with infinite constraints. We will show that for a fixed $\mathbf{x}$, the dual can be cast as a second-order cone program (SOCP) that can be solved tractably using commercial off-the-shelf solvers such as CVX. Hence, to solve (DRO-J), we can use grid search on $\mathcal{X}$ where, for each grid point $\mathbf{x}$, we solve an SOCP.

Next, we derive the SOCP formulation given $\mathbf{x}$. Note that $J^x(hx_e, x_f)$ is a piecewise-linear function in $h \in [h_\ell, h_u]$ with two breakpoints, $h_0 := (\lambda - \theta x_f)/x_e$ and $h_f := (\alpha x_f)/x_e$. To simplify our exposition, assume that $h_0 < h_f$ and that both breakpoints lie in $[h_\ell, h_u]$. (The technique can be easily adapted for other cases.) Hence, the dual problem is equivalent to the following:

$$\begin{aligned} \max_{t,s,r} \quad & t + \mu s + (\mu^2 + \sigma^2) r \\ \text{s.t.} \quad & t + hs + h^2 r \le a_1 + b_1 h, \quad \forall h \in [h_\ell, h_0], \\ & t + hs + h^2 r \le a_2 + b_2 h, \quad \forall h \in [h_0, h_f], \\ & t + hs + h^2 r \le a_3 + b_3 h, \quad \forall h \in [h_f, h_u] \end{aligned} \tag{8}$$

where a_i, b_i, c_i for $i = 1, 2, 3$ are appropriately defined from the functional form of $J^x(hx_e, x_f)$ in each of the subdomains of $[h_\ell, h_u]$.

Through a sequence of reformulations, we can cast each of the three constraint groups in (8) as a second-order cone constraint. We demonstrate this technique in the first constraint group. First, we apply the S-lemma (Polik and Terlaky 2007).

Lemma 1 **(S-lemma, Polik and Terlaky 2007)**. *Let* $q_a, q_b : \mathbb{R}^n \mapsto \mathbb{R}$ *be quadratic functions. Suppose there exists* $\bar{\mathbf{h}} \in \mathbb{R}^n$ *such that* $q_a(\bar{\mathbf{h}}) > 0$*. Then the implication* $[q_a(\mathbf{h}) \geq 0 \Longrightarrow q_b(\mathbf{h}) \geq 0]$ *holds if and only if there exists* $z \geq 0$ *such that* $q_b(\mathbf{h}) \geq z q_a(\mathbf{h})\ \forall \mathbf{h}$.

Define $q_a(h) := (h - h_\ell)(h_0 - h)$ and $q_b(h) := -rh^2 + (b_1 - s)h + (a_1 - t)$. Note that the first constraint group can be equivalently reformulated as the implication $[q_a(h) \geq 0 \Longrightarrow q_b(h) \geq 0]$. By S-lemma, this is equivalent to the constraint that, there exists $z_1 \geq 0$ such that $q_b(h) - z_1 q_a(h) \geq 0$ for all h, or

$$\begin{pmatrix} -r & \frac{1}{2}(b_1 - s) \\ \frac{1}{2}(b_1 - s) & a_1 - t \end{pmatrix} - z_1 \begin{pmatrix} -1 & \frac{1}{2}(h_\ell + h_0) \\ \frac{1}{2}(h_\ell + h_0) & -h_\ell h_0 \end{pmatrix} \succeq 0.$$

We can reformulate the above semidefinite constraint using the following equivalence:

Claim 1. $[4xy - z^2 \geq 0$ *and* $x \geq 0] \Longleftrightarrow [x + y \geq \sqrt{z^2 + (x-y)^2}]$

Hence, the first constraint group of (8) is equivalent to a second-order cone constraint:

$$-r + z_1 + a_1 - t + z_1 h_\ell h_0 \geq \left\| \begin{matrix} b_1 - s - z_1(h_\ell + h_0) \\ -r + z_1(1 - h_\ell h_0) - a_1 + t \end{matrix} \right\|_2,$$

where z_1 is a new non-negative decision variable. We can apply the sequence of reformulations to the remaining constraint groups. Hence, the dual program is equivalent to a second-order cone program (SOCP) with variables t, s, r and non-negative variables z_1, z_2, z_3.

Note that the S-lemma is critical in reformulating $J_{DRO}(\mathbf{x})$ into a SOCP. The S-lemma can only be used when u is piecewise-linear or piecewise-quadratic. However, we can approximate other forms of u (e.g., exponential) using a piecewise-linear function. Hence, we can use our technique even under a general function u to approximate the value of $J_{DRO}(\mathbf{x})$ by solving an SOCP.

■