Online Algorithms for Matching Platforms with Multi-Channel Traffic

Vahideh Manshadi
Yale School of Management, New Haven, CT, vahideh.manshadi@yale.edu

Scott Rodilitz
Stanford Graduate School of Business, Stanford, CA, scott.rodilitz@gmail.com

Daniela Saban
Stanford Graduate School of Business, Stanford, CA, dsaban@stanford.edu

Akshaya Suresh
Yale School of Management, New Haven, CT, akshaya.suresh@yale.edu

Two-sided platforms rely on their recommendation algorithms to help visitors successfully find a match. However, on platforms such as VolunteerMatch – which has facilitated millions of connections between volunteers and nonprofits – a sizable fraction of website traffic arrives directly to a nonprofit's volunteering page via an external link, thus bypassing the platform's recommendation algorithm. We study how such platforms should account for this *external traffic* in the design of their recommendation algorithms, given the goal of maximizing successful matches. We model the platform's problem as a special case of online matching, where (using VolunteerMatch terminology) volunteers arrive sequentially and probabilistically match with one opportunity, each of which has finite need for volunteers. In our framework, external traffic is interested only in their targeted opportunity; by contrast, *internal traffic* may be interested in many opportunities, and the platform's online algorithm selects which opportunity to recommend. In evaluating the performance of different algorithms, we refine the notion of competitive ratio by parameterizing it based on the amount of external traffic. After demonstrating the shortcomings of a commonly-used algorithm that is optimal in the absence of external traffic, we propose a new algorithm – Adaptive Capacity (`AC`) – which accounts for matches differently based on whether they originate from internal or external traffic. We provide a lower bound on `AC`'s competitive ratio that is increasing in the amount of external traffic and that is close to the parameterized upper bound we establish on the competitive ratio of any online algorithm. Our analysis utilizes a path-based, pseudo-rewards approach, which we further generalize to settings where the platform can recommend a ranked set of opportunities. Beyond our theoretical results, we demonstrate the strong performance of `AC` in a case study motivated by VolunteerMatch data.

Key words: matching platforms, online algorithms, competitive analysis, multi-channel traffic

1. Introduction

Online platforms have become increasingly prominent in facilitating social and economic connections in both the private and nonprofit sectors. In the private sector, the e-commerce platform Etsy has empowered over 2 million small-scale sellers to showcase their products to over 40 million buyers and has facilitated transactions on the scale of $4 billion.[1] In the nonprofit sector, the crowdfunding platform DonorsChoose has helped public school teachers to successfully solicit $314 million in donations for 1.7 million classroom projects.[2] Similarly, VolunteerMatch has enabled over 14 million connections between organizations and individuals looking for volunteering opportunities.

These platforms attract traffic through multiple channels. Some users organically visit the platform and rely on its recommendation algorithm to find a desired product or volunteering opportunity—we refer to these users as *internal traffic*. Other users, which we refer to as *external traffic*, follow an external direct link to a particular page. This external traffic is generated through a variety of off-platform outreach mechanisms, such as posting on social media or sending customized notifications. For example, an artist who sells their handmade products on Etsy may tweet about them, or an NGO may publicize their volunteering/donation opportunities on their Facebook page. In this paper, we aim to understand how these matching platforms can efficiently leverage traffic from *all* sources in order to maximize the number of successful transactions/connections.

This work is partly motivated by our collaboration with VolunteerMatch (VM), the largest nationwide platform that connects nonprofits with volunteers. More than 130,000 organizations—supporting a variety of social causes, ranging from human rights and literacy to helping seniors—have posted their volunteering opportunities on the VM website. Most of these organizations rely on volunteers who sign-up after visiting the VM website. Some of these organizations also generate sign-ups by publicizing their opportunities on other websites, such as LinkedIn or Facebook. Our analysis of VM data reveals two key facts. First, a significant portion of volunteer sign-ups come from external traffic: for example, 30% of all sign-ups made by NYC-based volunteers between August 1, 2020 and March 1, 2021 came from external traffic. Second, there is a significant disparity across opportunities in terms of both the total number of sign-ups and the source of those sign-ups.

To illustrate these two facts, in Figure 1 we plot the distribution of the number of sign-ups for a subset of opportunities that all requested 5 volunteer sign-ups.[3] Partitioning the sign-ups into two groups based on their source, we observe that the volume of sign-ups from external traffic (in purple) and from internal traffic (in green) varies substantially across opportunities.[4] From the platform's perspective, a key difference between external and internal traffic comes from whether or not the user's choice can be influenced: the platform cannot control the "landing page" for external traffic, but it can impact what internal traffic views (and thus the decisions made) via its recommendation

[1] https://www.sec.gov/Archives/edgar/data/1370637/000137063719000028/etsy1231201810k.htm

[2] https://www.donorschoose.org/about/impact.html

[3] This subset of 100 opportunities is a random sample of all virtual opportunities requesting 5 volunteer sign-ups between August 2020 and March 2021.

[4] We only observe the source for a subset of sign-ups, as described in Appendix C. We estimate the source of each opportunity's sign-ups proportionally, based on this subset.

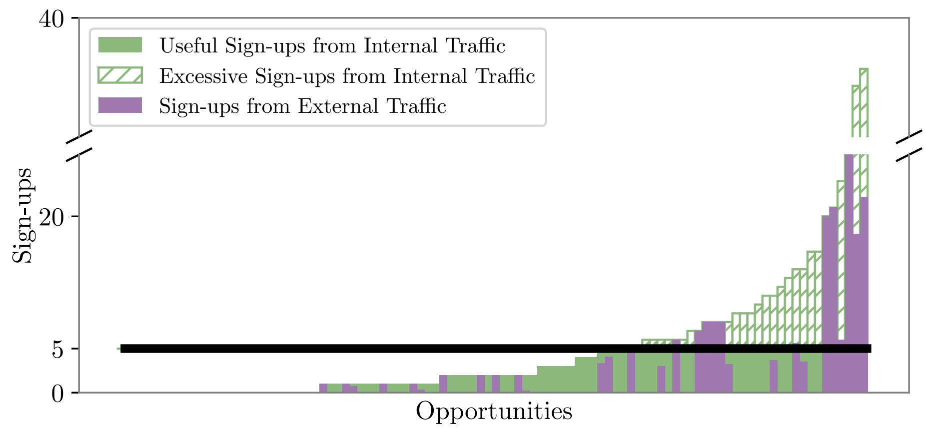


Figure 1 **Distribution of sign-ups on VM across a subset of opportunities requesting 5 volunteer sign-ups.**

algorithm. Through its search design, the platform can (potentially) re-distribute "excessive" sign-ups from internal traffic (i.e., sign-ups that exceed an opportunity's need) to opportunities with insufficient sign-ups, thereby helping VM achieve its strategic goal of maximizing the total number of "useful" sign-ups across opportunities.[5] For instance, for the subset of opportunities presented in Figure 1, in hindsight, 49% of sign-ups from internal traffic (the dashed green portions of the bars) could potentially have been re-directed to opportunities with insufficient sign-ups.

The above observations motivate our main research question: *how can matching platforms, such as VM, integrate external and internal traffic to maximize the number of useful sign-ups?* As the traffic pattern is generally unknown a priori and there is heterogeneity in the level of external traffic, making better *real-time* recommendations to internal traffic may be challenging.

1.1. Our Contributions

To study the above question, we introduce a framework for online matching with multi-channel traffic. Taking a competitive analysis approach, we show that existing algorithms—that are optimal in the absence of external traffic—fail to integrate such traffic efficiently; thus, we develop a new algorithm that effectively incorporates external traffic, resulting in near-optimal guarantees in certain regimes. Beyond worst-case guarantees, we illustrate the effectiveness of our algorithm in a simulation study calibrated on VM data. We describe each contribution in more detail next.

A model for online matching with multi-channel traffic: For concreteness, we utilize terminology from the context of VM and refer to the two sides of the matching platform as "opportunities" and "volunteers." In our setting, a fixed set of opportunities are posted on the platform,

[5] We note that the skewed sign-up distribution not only hurts opportunities with insufficient sign-ups, but it also harms other stakeholders. For instance, individuals that sign up for opportunities with excessive sign-ups may be discouraged if their attempts to volunteer are ignored or if they exert unnecessary effort. Additionally, organizations that receive excessive sign-ups may also incur/impose costs due to screening or training unnecessary volunteers.

each requiring a certain number of volunteers which we refer to as their "capacity." Volunteers arrive sequentially (in an arbitrary order) and are either external or internal traffic. External traffic directly views a specific opportunity's page and signs up with their *conversion probability* for that opportunity (i.e., the probability that the volunteer signs up for that opportunity conditional on viewing it). By contrast, internal traffic can be influenced by the platform's recommendation algorithm as follows: when an internal traffic volunteer arrives, the platform observes their conversion probabilities for each opportunity, and then must immediately and irrevocably recommend one such opportunity.[6] The goal of the platform is to maximize the total number of "useful" sign-ups, i.e., the total number of signups that don't exceed an opportunity's capacity.

In the absence of external traffic, the above problem can be viewed as an instance of the online bipartite B-matching problem with stochastic rewards and an adversarial arrival sequence. In this general framework, it has been shown that a simple myopic algorithm commonly-referred to as `MSVV` achieves the best-possible competitive ratio of $1-1/e$ (Mehta et al. 2007).[7] We augment this framework by modeling external traffic as arrivals with only one possible edge (e.g., volunteers that only consider one opportunity). The presence of external traffic reduces the complexity of making real-time decisions: the platform cannot change what external traffic volunteers will view, as they are only interested in one opportunity. Thus, in the extreme case where all capacity can be filled by external traffic, the platform trivially maximizes the number of useful sign-ups.

In light of the above observation, we parameterize problem instances based on the fraction of total capacity that can be filled by external traffic, which we call the *effective fraction of external traffic* (EFET), as formalized in Definition 2. For a given EFET, we define the competitive ratio of an algorithm to be the worst-case ratio between its outcome and that of a benchmark, among all instances with that EFET (see Definition 3). Our benchmark (denoted `OPT`) is a clairvoyant solution that a priori knows the sequence of arrivals as well as the sign-up realizations of external traffic, but only observes the sign-up realizations of internal traffic ex-post (see Definition 1). We study how the addition of external traffic improves the achievable competitive ratio.

Failure of channel-agnostic algorithms: To gain intuition, we first focus on a thought experiment where all of the external traffic arrives before any of the internal traffic. In such a setting, after the sign-ups from external traffic realize, the platform is faced with a standard instance of the online matching problem. Thus, by making recommendations in the *remaining* problem according

---

[6] In our base model (introduced in Section 3), we assume that the platform recommends a single opportunity. We consider a more general setting where the platform can present a ranking of opportunities in Section 5.

[7] Though Mehta et al. (2007) considers a setting with deterministic rewards, as noted in Mehta et al. (2013), the guarantee and the optimality of `MSVV` extend (asymptotically) to a B-matching setting with stochastic rewards when all capacities are sufficiently large. We will henceforth describe results only for the large-capacity setting; however, our technical results are all parameterized by the minimum capacity.

to an optimal algorithm like MSVV, we would hope to achieve a competitive ratio that is a convex combination of 1 and $1-1/e$. Indeed, in Proposition 1, we prove that this convex combination is an upper bound on any online algorithm. However, somewhat surprisingly, applying MSVV to the *entire* problem instance does not achieve this intuitive bound (Proposition 2). The suboptimality of this algorithm stems precisely from a lack of differentiation between external and internal traffic.

Adaptive Capacity (AC) algorithm: Building on the intuition developed in the thought experiment above, we introduce a new algorithm called *Adaptive Capacity* (AC) which reduces an opportunity's capacity by one whenever that opportunity receives a sign-up from external traffic. If all external traffic arrives before any internal traffic, AC achieves the upper bound in Proposition 2. However, in a general setting where external traffic can arrive at arbitrary times, AC does not have the information needed to reduce capacities up-front; instead, it *adaptively* reduces capacity after each sign-up from external traffic (see Algorithm 2).

Our main theoretical result establishes a lower bound (as a function of the EFET) on the competitive ratio of AC (see Theorem 2 and Figure 2b) parametrized by the *maximum conversion probability ratio* (MCPR), which we formally introduce in Definition 4. As the MCPR increases (e.g., in settings where volunteers have significant heterogeneity in their non-zero conversion probabilities), the AC algorithm's guarantee decreases. Fixing any MCPR, our lower bound curve starts at $1-1/e$ (when there is no external traffic) but weakly increases with the EFET and eventually breaks the barrier of $1-1/e$. To shed light on the limitations imposed by real-time decision making, we also establish an upper-bound (as a function of the EFET) on the competitive ratio of any online algorithm (Theorem 1). If the MCPR is 1 (which is the case, e.g., if each conversion probability is either 0 or 1), our upper bound nearly matches our lower bound on AC for any EFET. This is particularly intriguing because our algorithm does not know the volume of external traffic in advance; yet by adaptively reducing capacities, it achieves a near-optimal competitive ratio.

To analyze the competitive ratio of AC, we extend the LP-free approach in Goyal et al. (2020), which establishes a system of inequalities involving path-based "pseudo-rewards." To break the barrier of $1-1/e$ we leverage the observation that an algorithm cannot make a bad decision for external traffic, and thus we define pseudo-rewards based on the source of the traffic. Moreover, as the volume of external traffic varies across opportunities, we move beyond an opportunity-level analysis, and instead bound the "global" value of AC relative to OPT.

In Section 5, we extend our model to one where the platform recommends a ranked subset of opportunities and the volunteer signs up for (at most) one of these opportunities. We naturally generalize the AC algorithm to a ranking algorithm denoted AC-R (see Equation (12)). Using our flexible proof technique, we establish a lower bound on AC-R's competitive ratio for arbitrary volunteer choice functions (Proposition 4) as well as a stronger lower bound for a special class of volunteer choice functions (see Definition 5 and Proposition 5).

Case study based on VM: To illustrate the effectiveness of our algorithm in practice, we evaluate its performance on a problem instance constructed using a VM dataset that enables us to preserve real-life patterns of external traffic and heterogeneity in conversion probabilities. We show that our `AC` algorithm significantly outperforms a proxy for current practice on VM (Figure 4a). It achieves this level of performance by reducing the number of excessive sign-ups, thereby utilizing internal traffic more efficiently (Figure 5).

2. Related Work

Our work relates to and contributes to several streams of literature.

Generalized Online Matching: The rich literature on online matching started with the seminal work of Karp et al. (1990); given the scope of this literature, we discuss only a few papers and kindly refer the reader to Mehta et al. (2013) for a comprehensive survey. We model the platform's problem as a generalized instance of online B-matching (Kalyanasundaram and Pruhs 2000), which has been extensively studied in the context of online advertising (Mehta et al. 2007, Buchbinder et al. 2007, Balseiro et al. 2020, Udwani 2021).[8] Variants of online B-matching problems have been recently proposed to study a variety of problems arising in online platforms, including real-time assortment decisions (Golrezaei et al. 2014, Ma and Simchi-Levi 2020, Aouad and Saban 2020, Désir et al. 2021) and online allocation of reusable resources (Feng et al. 2019, Goyal et al. 2020, Rusmevichientong et al. 2020, Gong et al. 2021). We contribute to this line of work by introducing a variant of online matching motivated by platforms with multi-channel traffic.

In our model, each external traffic volunteer corresponds to a degree-one arriving node. Our `AC` algorithm effectively incorporates these degree-one nodes, and not only breaks the barrier of $1-1/e$ given a sufficient amount of external traffic, but also achieves a near-optimal competitive ratio in certain parameter regimes. In a similar vein, the work of Buchbinder et al. (2007) and Naor and Wajc (2018) impose a bound on the degree of *all* nodes in one or both sides and show that one can improve upon a competitive ratio of $1-1/e$ for such structured instances. We emphasize that our work differs from these papers, as we make no assumption on the degree of internal traffic. Our proof technique builds on the flexible LP-free approach of Goyal and Udwani (2019) and Goyal et al. (2020), which we use to distinguish between external and internal traffic in our analysis.

Hybrid Traffic Models: The challenge of integrating different channels of traffic arises in other application domains as well, such as retail and e-commerce. Dzyabura and Jagabathula (2018) study

[8] Our framework allows for stochastic rewards, which can introduce additional challenges (Mehta and Panigrahi 2012, Goyal and Udwani 2019). We sidestep this challenge by parameterizing our results based on the minimum capacity and by focusing on the large-capacity regime, following the approach of this literature.

a retail setting where the firm offers products through both offline and online channels. Consumers are a mixture of three types: those who visit only online or only offline, and those who visit the store before making a purchasing decision online (and thus their preference may be impacted by the products showcased in the offline store). They study assortment problems for this mixture of consumers. In the context of e-commerce, Esfandiari et al. (2015) and Hwang et al. (2021) consider online allocation problems where the traffic is composed of a stochastic (predictable) component as well as an adversarial (unpredictable) one. We contribute to this line of work by introducing a new hybrid traffic model that consists of external and internal traffic.

Design of Matching Platforms: Motivated by the rapid growth of online matching platforms, recent work has shed light on how platform design can influence matching outcomes, e.g., in the context of labor markets (Aouad and Saban 2020), crowdsourcing (Manshadi and Rodilitz 2022), affordable housing (Arnosti and Shi 2020), ridesharing (Besbes et al. 2021), and dating markets (Ríos et al. 2020). Among other insights, this line of research analyzes the relative merits of different pricing/compensation policies (Alaei et al. 2022, Elmachtoub et al. 2022), demonstrates the value of limiting user choice (Immorlica et al. 2021, Kanoria and Saban 2021), and provides guidance on which assortments to show users of two-sided platforms (Ashlagi et al. 2019, Aouad and Saban 2020, Feldman and Segev 2022). We add to the platform design literature by studying how online matching platforms should adjust their recommendations to account for external traffic.

3. Model

In this section, we first formally introduce our model for the problem that a platform faces when providing recommendations in the presence of multi-channel traffic, which is a variant of online matching. (For ease of exposition, we will use terminology from the context of a volunteer matching platform to describe the model.) We then describe the platform's objective and the metric of a competitive ratio, which we will use to evaluate any online algorithm.

Each problem instance $\mathcal{I}$ consists of a static set of opportunities on the platform (denoted $\mathcal{S}$), a finite horizon of T periods, and a sequence of T volunteers who arrive to the platform (denoted $\vec{\mathbf{A}}$). We index opportunities with i from $i=1$ to $n=|\mathcal{S}|$. Each opportunity i has *capacity* c_i, which represents the total number of volunteer sign-ups needed by opportunity i. In each period t, the t^{th} volunteer in sequence $\vec{\mathbf{A}}$ arrives to the platform. As each period corresponds a volunteer arrival, we index volunteers according to their arrival time, i.e., volunteer t arrives at time t for $t \in [T]$.[9]

Volunteer dynamics: When volunteer t arrives, the platform observes its type, which consists of two components. The first component of a volunteer's type is its source, either EXT or INT, which

[9] For any $n \in \mathbb{N}$, we use $[n]$ to denote the set $\{1, 2, \dots, n\}$.

indicates whether the volunteer arrives to the platform as external or internal traffic, respectively. This is our way of modeling the multi-channel nature of traffic to the platform. We use $\mathcal{V}^{\text{EXT}}$ (resp. $\mathcal{V}^{\text{INT}}$) to denote the set of volunteers who arrive as external traffic (resp. internal traffic).

The second component of a volunteer's type is a vector $\boldsymbol{\mu}_t = \{\mu_{i,t} : i \in \mathcal{S}\}$, where $\mu_{i,t}$ is the pair-specific *conversion probability* with which volunteer t will sign-up for opportunity i if the volunteer views opportunity i. As motivated in the introduction, we assume that whenever external traffic arrives, they cannot be influenced by the platform and instead directly view their targeted opportunity, denoted i_t^*. By contrast, the platform chooses the opportunity that internal traffic views (as formalized below). After viewing an opportunity and making a sign-up decision, the volunteer leaves the platform.

Platform's Decisions and Objective: Upon each arrival, the platform observes the volunteer's type, i.e., their source as well as their pair-specific conversion probabilities. The platform then must immediately and irrevocably recommend a single opportunity to volunteer t, denoted $S_t \in \mathcal{S} \cup \{0\}$.[10] (In Section 5, we discuss how our model and results generalize to settings where the platform provides a ranked set of recommendations.) For external traffic, even though the platform plays no role in the volunteer's decision, we adopt the convention that the platform recommends $S_t = i_t^*$. The platform's recommendation for internal traffic can depend on the current volunteer's type, opportunity capacities, and the full history of volunteer arrivals and decisions. The volunteer then (deterministically) views the recommended opportunity, and signs up according to their pair specific conversion probability. We use the random variable $\xi_t(S_t) \in \{S_t, 0\}$ to denote the volunteer's sign-up decision when presented with the recommendation S_t.

The platform's objective is to maximize the amount of capacity filled by all volunteers (either internal or external traffic). We assume that all the sign-ups for an opportunity beyond its capacity *provide no value*. In the context of volunteer matching, these "excessive" sign-ups represent an ineffective use of volunteers, but can also have significant negative side effects, such as overwhelming the volunteer-management staff for that opportunity due to costly screening and reducing volunteer engagement due to under-utilization (Sampson 2006). (In other contexts such as e-commerce, the platform may be naturally constrained based on capacities.)

In pursuit of this objective, the platform follows an online recommendation algorithm $\pi \in \Pi$. For a volunteer arriving at time t, let opportunity S_t^π denote the (possibly random) opportunity recommended by algorithm π. Then, the expected amount of filled capacity generated by π (henceforth referred to as the expected *value* of π) on instance $\mathcal{I}$ is given by

$$\pi(\mathcal{I}) \quad = \quad \mathbb{E}\left[\sum_{i \in \mathcal{S}} \min\left\{c_i, \sum_{t \in [T]} \mathbb{1}[\xi_t(S_t^\pi) = i]\right\}\right],$$

[10] We introduce a "dummy" opportunity with index 0, which we use to indicate when the platform does not recommend an opportunity and when a volunteer does not sign-up for an opportunity.

where the expectation is taken with respect to the volunteers' sign-up realizations and, possibly, the randomized decisions by the algorithm.

Performance metric: To assess the quality of any proposed online algorithm π, we compare its expected value to that of an optimal clairvoyant algorithm $\texttt{OPT}$ on the same instance, denoted by $\texttt{OPT}(\mathcal{I})$. Consistent with the literature, we assume that $\texttt{OPT}$ operates with *a priori* knowledge of the exact sequence of volunteer arrivals $\vec{\mathbf{A}}$. Moreover, to compare ourselves against a benchmark that utilizes external traffic to the fullest extent possible, we strengthen $\texttt{OPT}$ further by assuming it has foreknowledge of the sign-up decisions of all external traffic. We note that this results in a stronger $\texttt{OPT}$ than one that does not have foreknowledge of the sign-up decisions of any arrivals, and thus will understate the performance of π compared to an $\texttt{OPT}$ that has knowledge of only the arrival sequence $\vec{\mathbf{A}}$.[11] This stronger benchmark aids in our analysis, as will become clear in Section 4. We formalize our notion of the benchmark $\texttt{OPT}$ in the following definition.

Definition 1 (Optimal Clairvoyant Algorithm) *The optimal clairvoyant algorithm is the solution to a dynamic program (of exponential size) which takes as input the instance $\mathcal{I}$ as well as the sign-up decisions of all external traffic throughout the time horizon. Upon the arrival of each internal traffic volunteer, the optimal clairvoyant algorithm recommends an opportunity $S_t^{\texttt{OPT}} \in \mathcal{S} \cup \{0\}$ that maximizes the total expected amount of filled capacity, given the sign-up history up to that point and the inputs to the program. Whenever there is more than one opportunity in this set of optimal opportunities, we use the convention (without loss of optimality) that $\texttt{OPT}$ deterministically recommends the opportunity in this set with lowest index.*

We highlight that our definition of $\texttt{OPT}$ ensures that it fills as much capacity as possible with external traffic. To see this, first note that $\texttt{OPT}$ knows in advance how much capacity *can* be filled by external traffic. Furthermore, if capacity can be filled by external traffic, then $\texttt{OPT}$ will never fill it with internal traffic instead: our convention for breaking ties in favor of opportunities with the lowest index implies that $\texttt{OPT}$ will recommend opportunity 0 (i.e., no opportunity) rather than wasting the sign-up from external traffic that will realize later.

The value of an algorithm relative to that of $\texttt{OPT}$ can depend significantly on the amount of capacity that can be filled by external traffic. For instance, if external traffic fills the entire capacity of each opportunity with certainty, then we can easily design an algorithm that achieves the same value as $\texttt{OPT}$. In this case, it would not matter how internal traffic was allocated, since external traffic alone will suffice to fill all capacity. Based on this observation, our performance metric will

[11] Allowing $\texttt{OPT}$ foreknowledge of the sign-up decisions of all external traffic is equivalent to assuming $\texttt{OPT}$ faces an arrival sequence where all external traffic arrives first, followed by internal traffic (maintaining the original arrival order within each traffic source). We thoroughly analyze settings with such an arrival pattern in Section 4.1.

be a function of both the online algorithm π as well as the expected fraction of capacity which can be filled by external traffic, as formalized below.

Definition 2 (Effective Fraction of External Traffic) *For a fixed instance $\mathcal{I}$, the effective fraction of external traffic (EFET) is the expected fraction of capacity which can be filled by external traffic. We use β to denote the EFET, where*

$$\beta(\mathcal{I}) = \frac{\sum_{i\in\mathcal{S}} \mathbb{E}\left[\min\{c_i, \sum_{t\in\mathcal{V}^{\text{EXT}}} \mathbb{1}[\xi_t(i_t^*)=i]\}\right]}{\sum_{i\in\mathcal{S}} c_i}. \tag{1}$$

We emphasize that our definition for `OPT` ensures that it fills a β fraction of capacity with external traffic in expectation, as `OPT` will never fill that capacity with internal traffic instead. For a given $\beta\in[0,1]$, we let $\mathcal{I}_\beta$ be the set of all possible instances where the EFET is β. Having defined our benchmark `OPT` and the parameter β, we are ready to define our performance metric. We will evaluate the performance of any online algorithm via the competitive ratio parametrized by β.

Definition 3 (Competitive Ratio) *The competitive ratio of an algorithm π for any effective fraction of external traffic $\beta\in[0,1]$ is given by:*

$$\textit{CompRatio}(\pi,\beta) \quad = \quad \min_{\mathcal{I}\in\mathcal{I}_\beta} \frac{\pi(\mathcal{I})}{\texttt{OPT}(\mathcal{I})} \tag{2}$$

By taking the minimum value of this ratio over all instances in $\mathcal{I}_\beta$, the competitive ratio provides a guarantee against even an adversarially-chosen instance.

To conclude this section, we revisit the connection with the online matching problems discussed in Section 2. The competitive ratio is a standard metric in this literature (see, e.g., Mehta et al. 2007), though the competitive ratio is commonly taken with respect to *all* possible instances. (In our setting, the domain of all possible instances is equivalent to the union over domains $\mathcal{I}_\beta$ for all $\beta\in[0,1]$.) In this work, motivated by the nature of external traffic that constitutes a considerable portion of traffic on some matching platforms, we explore how imposing structure on the problem (in the form of the EFET β) impacts the achievable competitive ratio.

4. Results

We start by considering a special case where all external traffic arrives before any internal traffic in Section 4.1. This special case provides intuition behind the shortcomings of known algorithms and motivates the need for our Adaptive Capacity (`AC`) algorithm. Building on this intuition, in Section 4.2 we establish a family of lower bounds on the competitive ratio of `AC` in a general setting, and we upper bound the competitive ratio of any online algorithm. Section 4.3 elaborates on the intuition and implications of our results. Finally, Section 4.4 provides the proof sketch of our main result.

4.1. Warm-up: External Traffic Arrives First

Let us first consider a setting where the platform observes all the external traffic before the arrival of any internal traffic. Any recommendation algorithm would use the same amount of external traffic as `OPT`, as we follow the convention that the platform cannot influence the decision of external traffic. However, an online algorithm may make sub-optimal recommendations to internal traffic, as it does not know which opportunities can be filled by future volunteers and which opportunities cannot. In settings without external traffic, this leads to a "barrier" of $1-1/e$, which is achievable asymptotically as the minimum capacity $\underline{c} = \min_{i\in[n]} c_i$ tends to infinity (Mehta et al. 2007).[12] Building on this intuition, the following proposition establishes an upper bound on the competitive ratio of any online algorithm.

Proposition 1 (Upper Bound when All External Traffic Arrives First) *Suppose that all external traffic arrives before internal traffic. Then, for any effective fraction of external traffic β and any minimum capacity, no online algorithm can achieve a competitive ratio greater than $\beta + (1-\beta)(1-1/e)$.*

The proof of Proposition 1 (which is presented in Appendix A.1) adjusts the hard instance presented in Mehta et al. (2007) by appending external traffic at the beginning of the arrival sequence, such that the EFET is equal to β.

Based on Proposition 1, one may ask: is it possible to design an online algorithm that achieves this upper bound, at least asymptotically? Intuitively, the answer should be yes. As noted above, in the absence of external traffic, it is possible to design algorithms that asymptotically achieve a competitive ratio of $1-1/e$. Building on such results, we should be able to design an algorithm that first fills (on average) a β fraction of capacity with external traffic, and then – based on the capacities that remain – treats the internal traffic portion of the problem as a typical instance of online matching, for which we can achieve a $1-1/e$ fraction of the offline solution `OPT`. Overall, this would lead to an asymptotic competitive ratio of at least $\beta + (1-\beta)(1-\frac{1}{e})$, as desired. However, a naive approach that only relies on existing algorithms does not achieve such a competitive ratio.

4.1.1. The failure of `MSVV`. A prime candidate to achieve this level of performance is the well-known algorithm introduced in Mehta et al. (2007), commonly referred to as `MSVV`. This algorithm achieves, asymptotically, the best-possible competitive ratio of $1-1/e$ for our online matching problem in the absence of external traffic, i.e., when $\beta = 0$.

The idea behind the `MSVV` algorithm is very simple. To balance the trade-off between the upside of recommending the opportunity with the highest conversion probability and the downside of

[12] Henceforth, we use "asymptotically" to refer to the regime where $\underline{c} \to \infty$. Notably, a competitive ratio of $1-1/e$ is *not* attainable in the finite-capacity regime (Mehta and Panigrahi 2012).

Algorithm 1 MSVV Algorithm (Mehta et al. 2007)

Initialize $\texttt{MSVV}_{i,0} = 0$, $\mathrm{FR}^{\texttt{MSVV}}_{i,0} = 0$ for all $i \in [n]$.
for $t \in [T]$ **do**
  if volunteer $t \in \mathcal{V}^{\text{EXT}}$ **then**
    Recommend $S^{\texttt{MSVV}}_t = i^*_t$ (i.e., recommend the unique targeted opportunity).
  else
    Recommend $S^{\texttt{MSVV}}_t \in \operatorname{argmax}_{S\in[n]\cup\{0\}} \mu_{S,t} \cdot \psi(\mathrm{FR}^{\texttt{MSVV}}_{S,t-1})$, where ψ is defined in (3).
  for $i \in [n]$ **do**
    $\texttt{MSVV}_{i,t} = \min\{c_i, \texttt{MSVV}_{i,t-1} + \mathbb{1}[\xi_t(S^{\texttt{MSVV}}_t) = i]\}$; $\mathrm{FR}^{\texttt{MSVV}}_{i,t} = \texttt{MSVV}_{i,t}/c_i$

reaching an opportunity's capacity before the end of the horizon, MSVV weighs each opportunity's conversion probability with the following decreasing trade-off function of the opportunity's fill rate:

$$\psi(\mathrm{FR}) = 1 - \exp(\mathrm{FR} - 1). \tag{3}$$

Opportunity i's fill rate under MSVV after the arrival of volunteer t (denoted $\mathrm{FR}^{\texttt{MSVV}}_{i,t}$) is the fraction of opportunity i's capacity (c_i) that is filled at that time. We formally present MSVV in Algorithm 1.[13]

Surprisingly, MSVV does not achieve the desired competitive ratio of $\beta + (1-\beta)(1-\frac{1}{e})$ in the setting where all external traffic comes first, as established by the following proposition.

Proposition 2 (Upper Bound on MSVV when All External Traffic Arrives First) *Suppose external traffic arrives before internal traffic. Then for any effective fraction of external traffic β and any minimum capacity, the competitive ratio of* MSVV *is at most*

$$1 - \frac{1-\hat{\alpha}_1}{\exp\left(\exp(-\hat{\alpha}_1/(1-\hat{\alpha}_1))\right)} \tag{4}$$

where $\hat{\alpha}_1$ is the unique solution in $[0,1]$ to $\beta = \hat{\alpha}_1 + (1-\hat{\alpha}_1)\Big(\exp\big(-\hat{\alpha}_1/(1-\hat{\alpha}_1)\big) - 1\Big)$.

In Figure 2a, we illustrate the upper bound on the competitive ratio of MSVV given by (4). There is a significant gap between the upper bound on the competitive ratio of MSVV (dashed red curve) and the potentially-achievable frontier characterized in Proposition 1 (solid blue line). The shortcomings of MSVV stem from its definition of an opportunity's fill rate, i.e., $\mathrm{FR}^{\texttt{MSVV}}_{i,t} = \texttt{MSVV}_{i,t}/c_i$, which accounts for internal and external traffic in an identical fashion. Under MSVV, the opportunities that receive sign-ups from external traffic will have strictly positive fill rates when internal traffic arrives, and thus will be de-prioritized. The proof of Proposition 2 (presented in Appendix A.2) builds on this intuition: we design a family of instances in which MSVV (sub-optimally) withholds internal

[13] If there are multiple recommendations that satisfy MSVV's optimality criteria, we follow the convention of recommending the one with the lowest index.

traffic from opportunities that initially receive external traffic. In these instances, for $\beta \in (0,1)$, the amount of capacity filled by internal traffic under MSVV is less than a $1-1/e$ factor of the amount of capacity filled by internal traffic under OPT. Consequently, it would appear that in order to achieve a competitive ratio of $\beta+(1-\beta)(1-\frac{1}{e})$, we must design an algorithm that incorporates the source of traffic into its decision-making. To that end, we next introduce our *Adaptive Capacity* (AC) algorithm, which accounts for the amount of filled capacity separately based on source.

4.1.2. Accounting for the source of traffic: the Adaptive Capacity algorithm. Similar to MSVV, the AC algorithm uses the exponential trade-off function ψ, as defined in (3), and it recommends the opportunity with the greatest weighted conversion probability, i.e., the opportunity i that maximizes $\mu_{i,t}\cdot\psi(\mathrm{FR}_{i,t-1})$.[14] However, AC *crucially differs from* MSVV *in its definition of an opportunity's fill rate.* The fill rate definition used by MSVV aggregates all sign-ups in the numerator; that is, it defines an opportunity's fill rate as $\mathrm{FR}^{\mathtt{MSVV}}_{i,t}=(\mathtt{MSVV}_{i,t})/c_i$. By contrast, AC aggregates sign-ups separately based on source, using counters $\mathtt{AC}^{\mathrm{EXT}}_{i,t}$ and $\mathtt{AC}^{\mathrm{INT}}_{i,t}$. It then removes any external traffic sign-ups from the total capacity (the denominator), i.e., $\mathrm{FR}_{i,t}=\mathtt{AC}^{\mathrm{INT}}_{i,t}/\left(c_i-\mathtt{AC}^{\mathrm{EXT}}_{i,t}\right)$. In other words, every time capacity is filled by external traffic, we *adaptively* reduce the capacity of that opportunity by one. We formally describe AC in Algorithm 2.

Algorithm 2 AC Algorithm

Initialize $\mathtt{AC}^{\mathrm{EXT}}_{i,0}=0$, $\mathtt{AC}^{\mathrm{INT}}_{i,0}=0$, and $\mathrm{FR}_{i,0}=0$ for all i in $[n]$.
for t in $[T]$ **do**
  if volunteer t in $\mathcal{V}^{\mathrm{EXT}}$ **then**
    Recommend $S^{\mathtt{AC}}_t:=j$, where $j=i^*_t$ (i.e., recommend the unique targeted opportunity).
    $\mathtt{AC}^{\mathrm{EXT}}_{j,t}=\min\{c_j-\mathtt{AC}^{\mathrm{INT}}_{j,t},\mathtt{AC}^{\mathrm{EXT}}_{j,t-1}+\mathbb{1}[\xi_t(j)=j]\}$; $\mathtt{AC}^{\mathrm{INT}}_{j,t}=\mathtt{AC}^{\mathrm{INT}}_{j,t-1}$
  else
    Recommend $S^{\mathtt{AC}}_t:=j$, where $j\in\mathrm{argmax}_{S\in[n]\cup\{0\}}\mu_{S,t}\cdot\psi(\mathrm{FR}_{S,t-1})$, where ψ is defined in (3).
    $\mathtt{AC}^{\mathrm{INT}}_{j,t}=\min\{c_j-\mathtt{AC}^{\mathrm{EXT}}_{j,t},\mathtt{AC}^{\mathrm{INT}}_{j,t-1}+\mathbb{1}[\xi_t(j)=j]\}$; $\mathtt{AC}^{\mathrm{EXT}}_{j,t}=\mathtt{AC}^{\mathrm{EXT}}_{j,t-1}$
  $\mathrm{FR}_{j,t}=\mathtt{AC}^{\mathrm{INT}}_{j,t}/(c_j-\mathtt{AC}^{\mathrm{EXT}}_{j,t})$
  for i in $[n]\setminus\{j\}$ **do**
    $\mathtt{AC}^{\mathrm{EXT}}_{i,t}=\mathtt{AC}^{\mathrm{EXT}}_{i,t-1}$; $\mathtt{AC}^{\mathrm{INT}}_{i,t}=\mathtt{AC}^{\mathrm{INT}}_{i,t-1}$; $\mathrm{FR}_{i,t}=\mathrm{FR}_{i,t-t}$

In the following, we establish that the competitive ratio of AC is asymptotically optimal when external traffic arrives before internal traffic. Intuitively, in this warm-up setting, AC implements the solution discussed in the beginning of this section: it reduces capacities based on the number of

[14] If there are multiple recommendations that satisfy AC's optimality criteria, we follow the convention of recommending the one with the lowest index.

sign-ups from external traffic and then, for internal traffic, it runs `MSVV` on the *remaining* capacities. Building on this intuition, the following proposition lower-bounds the competitive ratio of `AC`.

Proposition 3 (Lower Bound on `AC` when All External Traffic Arrives First) *Suppose all external traffic arrives before internal traffic. Then for any effective fraction of external traffic β and any minimum capacity $\underline{c}$, the competitive ratio of* `AC` *is at least $\beta + (1-\beta)(1-1/e) - \underline{c}^{-1}$.*

The lower bound given in Proposition 3 (which we prove in Appendix A.3) asymptotically achieves the upper bound established in Proposition 1 (shown by Figure 2a).

To conclude this section, we note that even though this warm-up setting is unrealistic and studied solely to develop intuition, it is roughly equivalent to the more realistic setting where the arrival order of the external and internal traffic can be arbitrarily mixed, but the amount and type of external traffic can be accurately predicted in advance. In such a setting, a variant of `AC` that reduces capacities by the *expected* amount of sign-ups from external traffic and then, for each internal traffic arrival, runs `MSVV` on the remaining capacities will be asymptotically optimal. Intuitively, this is akin to the `AC` algorithm in the warm-up setting, which reduces capacities by the *realized* amount of sign-ups from external traffic. The difference between the expected and the realized amount of capacity filled by external traffic has a vanishing impact on the competitive ratio in the asymptotic regime. This more realistic setting illustrates the need to differentiate internal and external traffic beyond the warm-up case and motivates us to analyze `AC` in an even more general setting.

4.2. Performance of the `AC` Algorithm Under More General Arrivals

The previous section focused on a setting where external traffic arrives to the platform first, and we observed that the competitive ratio of `AC` significantly improves upon the fundamental barrier of $1-1/e$ (which we remind is the upper-bound in the absence of external traffic). We now investigate the competitive ratio of `AC` when the arrival sequence of volunteer types is completely unknown. To that end, we will compare the competitive ratio of `AC` to an upper bound on the competitive ratio of any online algorithm, as a function of the EFET β.

In contrast with the setting previously described, the `AC` algorithm cannot always observe the sign-ups from external traffic before making recommendations for internal traffic. As a consequence, when internal traffic arrives, the `AC` algorithm may inadvertently recommend an opportunity which could be filled entirely by later-arriving external traffic. This is not only a limitation of the `AC` algorithm: no online algorithm has access to information about future external traffic. However, the information available to `OPT` is unchanged: it still has *a priori* knowledge of the amount of capacity that can be filled by external traffic. We should intuitively expect the achievable competitive ratio will decrease in this setting (compared to the warm-up setting), as one could construct hard examples where valuable information about external traffic is not revealed until the end of the arrival sequence (e.g., if all external traffic arrives after all internal traffic).[15] Building on this intuition, we modify the hard instance of Mehta et al. (2007) by replacing the tail end of the arrival sequence with carefully-designed external traffic. This modification allows us to establish the following family of upper bounds on the competitive ratio of any online algorithm.

Theorem 1 (Upper Bound on Competitive Ratio) *For any effective fraction of external traffic β and any minimum capacity, no online algorithm can achieve a competitive ratio better than* $\max\{1-1/e, 1+\beta\log(\beta)\}$.

We illustrate the upper bound as a function of the EFET β in Figure 2b (blue curve), and we formally prove this result in Appendix A.4. Naturally, one wonders whether the `AC` algorithm can come close to attaining this upper bound. We find that the answer depends in part on the maximum conversion probability ratio, a quantity we formally define below.

Definition 4 (Maximum Conversion Probability Ratio) *For each volunteer t, let $\mathcal{S}_t$ denote the subset of opportunities i for which $\mu_{i,t} > 0$.*[16] *The* conversion probability ratio *(CPR) for volunteer t is given by $\frac{\max_{i\in\mathcal{S}_t}\mu_{i,t}}{\min_{i\in\mathcal{S}_t}\mu_{i,t}}$. The* maximum conversion probability ratio *(MCPR), denoted by σ, is the maximum CPR across all volunteers, i.e.*

$$\sigma = \max_{t\in[T]}\left(\frac{\max_{i\in\mathcal{S}_t}\mu_{i,t}}{\min_{i\in\mathcal{S}_t}\mu_{i,t}}\right) \tag{5}$$

Before providing intuition behind the dependence on the MCPR σ, we first present the main result of this section, which is a family of lower bounds on the competitive ratio of the `AC` algorithm. These bounds are parameterized by the EFET β, the minimum capacity $\underline{c}$, and the MCPR σ.

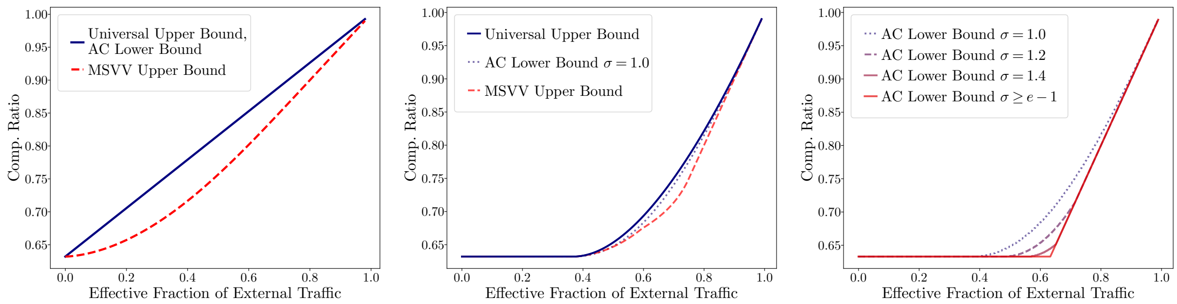
Theorem 2 (Lower Bound on `AC`'s Competitive Ratio) *Let the smallest capacity be given by $\underline{c}$ and let the maximum conversion probability ratio be at most σ. Then, for any effective fraction of external traffic β, the competitive ratio of the `AC` algorithm defined in Algorithm 2 (with ψ as defined in Eq. (3)) is at least $f(\beta,\underline{c},\sigma) = \max\{\beta, z^*\}$, where*

$$\begin{aligned} z^* \quad = \quad & \min_{z\in[0,1]} \; z \\ & \text{subject to} \quad z \;\geq\; e^{-1/\underline{c}}(1-1/e) \\ & \qquad\qquad\quad\;\; z \;\geq\; e^{-1/\underline{c}}\breve{g}\left(\max\{0,\beta-\sigma+z\}, z-\max\{0,\beta-\sigma+z\}\right), \end{aligned} \tag{6}$$

---

[15] We remark that even though many hard instances involve all external traffic arriving after all internal traffic, the two algorithms that we consider (i.e., `AC` and `MSVV`) do not exhibit performance that is monotonic in the arrival order of external traffic vis-à-vis internal traffic.

[16] Without loss of generality, we assume that for all volunteers, there is at least one opportunity for which they have a strictly positive conversion probability. Otherwise, we can simply remove that volunteer and re-index.



(a) (b) (c)

Figure 2 In the asymptotic regime, we present: a lower bound on the CR of AC and upper bounds on the CR of MSVV and any online algorithm (a) when all external traffic arrives first (Propositions 1, 2, and 3), and (b) under general arrivals (Theorems 1 and 2 and Proposition 6 in Appendix A.5). (c) The lower bound on the competitive ratio of AC for various values of σ under general arrivals.

and $\breve{g}(x_1,x_2)$ *denotes the lower convex envelope of* $g(x_1,x_2)$ *over the domain* $\mathcal{D}=\{(x_1,x_2)\in\mathbb{R}^2_{\geq 0}: x_1+x_2\leq 1\}$,[17] *where*

$$g(x_1,x_2)=1-\frac{1}{e}+x_1+(1-x_1)\psi\left(\frac{x_2}{1-x_1}\right)-\psi\left(x_2\right). \tag{7}$$

We next discuss the key takeaways from our main results; we defer a proof sketch of Theorem 2 to Section 4.4, with the remaining details provided in Appendix A.6.

4.3. Discussion of Results

Throughout this section, we restrict our attention to the asymptotic regime where $\underline{c}$ approaches infinity.[18] We begin to unpack Theorem 2 by first focusing on instances where $\sigma=1$. This class of instances includes the standard online B-matching problem (Kalyanasundaram and Pruhs 2000), where each volunteer has deterministic binary compatibility with an arbitrary subset of opportunities. In Figure 2b, we plot the lower bound on the competitive ratio of the AC algorithm when $\sigma=1$ (dotted purple line) in comparison to the upper bound on the competitive ratio of any online algorithm (solid blue line).[19] There is minimal gap between these lower and upper bounds, which suggests that the guarantee provided by our AC algorithm is close to the best-possible one for any EFET $\beta\in[0,1]$. This is particularly intriguing because our AC algorithm does not need to know the value of β in order to achieve a near-optimal guarantee for that EFET (in settings where $\sigma=1$).

Although we have demonstrated the superior competitive ratio of AC compared to MSVV when external traffic comes first (Propositions 2 and 3), it is natural to wonder if AC outperforms MSVV

[17] The lower convex envelope of a function g over a domain $\mathcal{D}$ is the supremum of all convex functions that are less than or equal to g on domain $\mathcal{D}$.

[18] As evident from (6), the value of $\underline{c}$ only impacts our lower bound via $e^{-1/\underline{c}}$, which is similar to the dependence on minimum capacity found in prior literature (see, e.g., Buchbinder et al. 2007, Goyal et al. 2020).

[19] We note that the upper bound holds for any minimum capacity $\underline{c}$ as well as any MCPR σ.

in the general case. In Proposition 6 (presented in Appendix A.5), we rule out this possibility by providing an upper bound on the competitive ratio of MSVV (shown by the dashed red curve in Figure 2b). The upper bound on MSVV is strictly below our lower bound on AC (by a multiplicative factor up to 3.9%) for $\beta \in [0.48, 0.94]$. Outside of that range, the upper bound on MSVV and the lower bound on AC are essentially indistinguishable (the two differ by at most 0.1%). This comparative analysis suggests that in the regime where $\sigma = 1$, differentiating between internal and external traffic (by using the AC algorithm) leads to an improved competitive ratio.

Maintaining (for now) our focus on settings where $\sigma = 1$, we next aim to better understand the relationship between the EFET β and our lower bound on the competitive ratio of AC. Similar to our tight bound in the setting where external traffic arrives first (as given by Propositions 1 and 3 in Section 4.1), the lower bound on the competitive ratio of the AC algorithm is non-decreasing in β. However, in the previous setting, the competitive ratio was linearly increasing in β. In contrast, in this general setting, no online algorithm can break the barrier of $1-1/e$ unless β exceeds $\beta^* = 1/e$.

As the dependence on e might suggest, there is a nice relationship between the fundamental barrier of $1-1/e$ (which we remind is the upper-bound in the absence of external traffic) and the threshold β^* on the EFET, as we next explain. Whenever AC generates a sign-up from external traffic, we know that OPT could not have made a "better" decision because external traffic (by definition) targets that particular opportunity. By leveraging the value of AC's "correct" decisions, we can demonstrate that AC has a competitive ratio strictly above $1-1/e$ *if* it fills a strictly positive amount of capacity with external traffic. Unfortunately, when the EFET is less than β^*, we cannot guarantee that AC fills *any* capacity with external traffic.

To see why, consider the following informal argument: Suppose volunteers have conversion probabilities of either 0 or 1, and suppose OPT allocates all volunteers and exactly fills all capacities. Even though AC attains the best-possible competitive ratio of $1-1/e$ in the absence of external traffic, there exists at least one instance where it "wastes" a $\beta^* = 1/e$ fraction of volunteers (i.e., AC *cannot* fill capacity with those volunteers). For any EFET $\beta \leq \beta^*$, we can construct a nearly-identical instance where the set of "wasted" volunteers includes *all* the external traffic. Indeed, under the AC algorithm, all external traffic is wasted on the instances which establish the upper bound of Theorem 1 for $\beta \in [0, \beta^*]$. However, when the EFET β strictly exceeds β^*, AC must fill a strictly positive amount of capacity with external traffic, which enables us to prove that AC's competitive ratio breaks the $1-1/e$ barrier (as we elaborate on in Section 4.4).

The informal argument of the prior paragraph falls apart, however, when applied to settings where volunteers may have different conversion probabilities for different (compatible) opportunities (i.e., where the MCPR $\sigma > 1$). In such settings, achieving a competitive ratio of $1-1/e$ is no longer a sufficient condition to ensure that the fraction of un-allocated volunteers is at most β^*.

Even when $\beta > \beta^*$, the fraction of un-allocated volunteers may include all the external traffic. As a consequence, we can no longer guarantee that the `AC` algorithm will break the $1-1/e$ barrier for every EFET greater than $\beta^* = 1/e$. We illustrate this challenge with the following example:

Example 1 (Limitation of `AC` for Unbounded σ) *Consider an instance with two opportunities (1 and 2) with capacities $c_1 = N$ and $c_2 = \frac{1}{e-1}N$ for sufficiently large N. There are $2N$ volunteers, and the first N volunteers are internal traffic with conversion probabilities given by*

$$\mu_{1,t} = 1, \qquad \mu_{2,t} = \frac{1-\exp\left(\frac{t-1}{N}-1\right)}{1-\exp(-1)} - \frac{1}{2N}.$$

The remaining N volunteers are external traffic for opportunity 1 with conversion probabilities of 1.

In Example 1, the EFET $\beta = 1-1/e$, as the capacity of opportunity 1 can be entirely filled with external traffic. The minimum capacity $\underline{c}$ and the MCPR σ are both arbitrarily large. (To see the latter, note $\mu_{1,N} = 1$ while $\mu_{2,N} = o(1)$.)[20] In this instance, `OPT` will recommend opportunity 2 to all internal traffic, and in expectation opportunity 2 will receive $\frac{1}{e-1}N - o(N)$ sign-ups. Then, external traffic arrives and fills opportunity 1, which means the amount of filled capacity under `OPT` is $\frac{e}{e-1}N - o(N)$.

In sharp contrast, `AC` will recommend opportunity 1 to all internal traffic volunteers, because the conversion probabilities in Example 1 are constructed such that $\mu_{1,t}\psi(\mathrm{FR}_{1,t}) > \mu_{2,t}\psi(0)$ for all $t \in [N]$. These internal traffic volunteers completely fill opportunity 1. Consequently, *no capacity is filled by external traffic under* `AC`, even though the EFET is $1-1/e$. In total, the amount of filled capacity under `AC` is N. Thus, in this example, the ratio between the expected value of `AC` and the expected value of `OPT` approaches $1-1/e$, despite the fact that the EFET $\beta = 1-1/e$.

Example 1 demonstrates that our analysis of the `AC` algorithm is tight for this set of parameters: it establishes an upper bound on the competitive ratio of the `AC` algorithm that matches the lower bound of Theorem 2 when $\beta = 1-1/e$ and in the limit as both $\underline{c}$ and σ approach infinity. Furthermore, this example sheds light on the challenges that arise when the MCPR exceeds 1. In line with this intuition, in Figure 2c, we show that `AC`'s guarantee is decreasing in the MCPR σ. Having discussed the comparative statics of our main result with respect to the EFET β and the MCPR σ, we now provide an overview of our proof technique.

[20] For two functions $d, l : \mathbb{N} \to \mathbb{R}$, $l(n) = o(d(n))$ if $\lim_{n\to\infty} \frac{l(n)}{d(n)} = 0$.

4.4. Proof Sketch of Theorem 2.

In this section, we present the proof sketch of Theorem 2. We start by noting that the lower bound on the competitive ratio of the AC algorithm, $f(\beta, \underline{c}, \sigma)$, is the maximum of two terms, meaning that each term represents a lower bound on the competitive ratio. To formally establish Theorem 2, we prove that the first term, β, is a lower bound on the competitive ratio (Lemma 1); then, we prove that the second term, z^*, is a lower bound on the competitive ratio (Lemma 2). The latter proof is more involved, and requires the introduction of path-based *pseudo-rewards*.

Lemma 1 (Lower Bound of β on $f(\beta, \underline{c}, \sigma)$) *Let the smallest capacity be given by $\underline{c}$ and let the maximum conversion probability ratio be at most σ. Then, for any effective fraction of external traffic β, the competitive ratio of the* AC *algorithm defined in Algorithm 2 (with ψ as defined in* (3)*) is at least β.*

Proof: The proof of Lemma 1 is immediate: we simply note that AC always recommends the targeted opportunity to external traffic. Applying the definition of the EFET (Definition 2) then ensures that at least a β fraction of capacity is filled in expectation. □

Lemma 2 (Lower Bound of z^* on $f(\beta, \underline{c}, \sigma)$) *Let the smallest capacity be given by $\underline{c}$ and let the maximum conversion probability ratio be at most σ. Then, for any effective fraction of external traffic β, the competitive ratio of the* AC *algorithm defined in Algorithm 2 (with ψ as defined in* (3)*) is at least z^* (with z^* as defined in* (6)*).*

Proof: The proof of Lemma 2 is fairly intricate, and our analysis leverages the LP-free approach developed in Goyal and Udwani (2019) and Goyal et al. (2020). This approach has proven useful in accounting for *post-allocation* stochasticity, e.g., stochastic rewards (as in Goyal and Udwani 2019) or stochastic usage duration (as in Goyal et al. 2020). In our setting with multi-channel traffic, we modify the approach to separately account for sign-ups based on their source, as the (potentially stochastic) amount of sign-ups from external traffic crucially impacts the guarantee that can be provided by the AC algorithm.

Central to this approach is the concept of *path-based pseudo-rewards*, i.e., values that are defined so as to keep track of the rewards that accrue during a particular run of an online algorithm relative to OPT. It is important to highlight that pseudo-rewards are defined purely for accounting purposes; in other words, they are not necessarily equivalent to the rewards of the algorithm on that particular run. (Nor are the pseudo-rewards equivalent to the dual solution of the underlying linear program, which is another commonly-used approach in the literature. See, e.g., Buchbinder et al. 2009.) These pseudo-rewards assist in the comparison between the online algorithm and OPT and ultimately allow us to establish a lower bound of z^* on the competitive ratio.

We divide the proof of Lemma 2 into three steps. In Step (1), we define appropriate pseudo-rewards for our setting. Our construction of pseudo-rewards departs from the approach of Goyal et al. (2020), as we define pseudo-rewards that are source-dependent. In Step (2), we use these pseudo-rewards to establish a lower bound on the expected value of AC that depends (in part) on the expected value of OPT (Lemmas 3 and 4). In contrast to the approach taken in Goyal et al. (2020), we cannot formulate a lower bound on the pseudo-rewards for each opportunity, as the amount of external traffic can be heterogeneous across opportunities. Instead, our more complex lower bound (on the expected sum of *all* pseudo-rewards) eventually enables us to break the competitive ratio barrier of $1-1/e$, but doing so requires an additional step. In this final step, Step (3), we construct a factor-revealing mathematical program (see Table 1) based, in part, on the lemmas of the previous step. Through analysis of this program, we place a lower bound of z^* on the competitive ratio of the AC algorithm (Lemmas 5 and 6).

Step 1: Defining Pseudo-Rewards

We begin by fixing a problem instance $\mathcal{I}$. We then define a *sample path* $\boldsymbol{\omega} = \{\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_T\}$, as the realizations of random variables that govern volunteer choices in this instance.[21] Formally, we interpret $\boldsymbol{\omega}_t$ as a vector of length n, where the i^{th} component of $\boldsymbol{\omega}_t$ (denoted $\omega_{i,t}$) indicates volunteer t's sign-up decision if the platform were to recommend opportunity i.[22] For the fixed instance $\mathcal{I}$ and for any fixed sample path $\boldsymbol{\omega}$, we will define pseudo-rewards $L_t(\mathcal{I}, \boldsymbol{\omega})$ for each volunteer $t \in [T]$, along with pseudo-rewards $K_i(\mathcal{I}, \boldsymbol{\omega})$ for each opportunity $i \in [n]$. Henceforth, to ease exposition, we suppress the dependence on the instance and the sample path.

Our pseudo-rewards L_t and K_i will depend on an opportunity's fill rate under AC along this fixed sample path, i.e., $\text{FR}_{i,t} = \frac{\texttt{AC}^{\text{INT}}_{i,t}}{c_i - \texttt{AC}^{\text{EXT}}_{i,t}}$, as well as on the realizations of volunteers' sign-up decisions under both AC (denoted $\xi_t(S_t^{\texttt{AC}})$) and OPT (denoted $\xi_t(S_t^{\texttt{OPT}})$).[23] Recall our convention that any algorithm (including AC) always recommends the targeted opportunity to external traffic. To ensure that we do not count sign-ups that exceed the capacity of an opportunity, we define $\tilde{\xi}_t(S_t^{\texttt{AC}})$ as the opportunity that volunteer t *fills capacity of* under AC. To be precise, if opportunity $\xi_t(S_t^{\texttt{AC}})$ has remaining capacity at time t, then $\tilde{\xi}_t(S_t^{\texttt{AC}}) = \xi_t(S_t^{\texttt{AC}})$; otherwise, $\tilde{\xi}_t(S_t^{\texttt{AC}}) = 0$.

Moreover, for this fixed instance $\mathcal{I}$ and along this fixed sample path $\boldsymbol{\omega}$, let $\mathcal{V}^0$ represent the subset of internal traffic for which OPT recommends opportunity 0, i.e., OPT does not recommend

[21] Fixing a set of realizations $\boldsymbol{\omega}$, the path of *any* deterministic algorithm (such as the AC algorithm) is uniquely determined. Hence, we refer to $\boldsymbol{\omega}$ as a sample path. That said, we emphasize that these realizations determine *all* possible choices for volunteers, not just the choices along the resulting sample path (i.e., the choices that result from the recommendations made by an algorithm).

[22] If the platform recommends opportunity 0, then the volunteer deterministically does not view (or sign up for) any opportunity.

[23] As noted above, we are suppressing these variables' dependence on the instance and the sample path. We emphasize that for a fixed instance and sample path, these variables are all deterministic.

any opportunity.[24] Based on our convention for $\texttt{OPT}$ introduced in Definition 1, a volunteer is in $\mathcal{V}^0$ if and only if all compatible opportunities have already reached their capacity for internal traffic. (Recall that $\texttt{OPT}$ knows *a priori* how much capacity will be filled by external traffic as it knows the realizations of those volunteers' sign-up decisions. This capacity is effectively reserved for external traffic, and only the remaining capacity will be filled by internal traffic. See Definition 1.)

With the above definitions, we are now ready to define the pseudo-rewards L_t and K_i.

$$L_t = \begin{cases} \sum_{i\in[n]} \psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}}) = i], & t \in \mathcal{V}^{\mathrm{EXT}} \cup \mathcal{V}^0 \\ \sum_{i\in[n]} \psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\xi_t(S_t^{\texttt{OPT}}) = i], & t \in \mathcal{V}^{\mathrm{INT}} \setminus \mathcal{V}^0 \end{cases} \tag{8}$$

$$K_i = \sum_{t\in[T]} (1 - \psi(\mathrm{FR}_{i,t-1}))\, \mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}}) = i] \tag{9}$$

For intuition behind our design of the volunteers' pseudo-rewards (i.e., the two cases in (8)), recall that our goal is to bound the difference between the values of $\texttt{AC}$ and $\texttt{OPT}$, which depends on the number of times $\texttt{OPT}$ makes a "better" recommendation than $\texttt{AC}$. Whenever external traffic arrives, $\texttt{OPT}$ will recommend the targeted opportunity, which cannot be better than the recommendation made by $\texttt{AC}$. Similarly, for internal traffic where $\texttt{OPT}$ does not recommend an opportunity (i.e., for $t \in \mathcal{V}^0$), then the recommendation made by $\texttt{OPT}$ cannot be better, in the sense that the objective is (weakly) increasing in the total number of sign-ups. In contrast, when internal traffic arrives and $\texttt{OPT}$ does make a recommendation, then this recommendation can be "better" than the recommendation made by $\texttt{AC}$. Hence, we define different pseudo-rewards for these arriving volunteers.

Step 2: Lower-bounding the Value of $\texttt{AC}$

This step of the proof involves two lemmas. First, in Lemma 3, we use the optimality criteria for the recommendations provided by the $\texttt{AC}$ algorithm to show that the expected sum of the L_t and K_i pseudo-rewards is a lower bound on the expected value of $\texttt{AC}$. (We use $\texttt{AC}$ to denote the value of the $\texttt{AC}$ algorithm along a fixed sample path for a fixed instance, again suppressing the dependence for ease of exposition.) Then, in Lemma 4, we use properties of the function ψ (as defined in (3)) to lower bound the expected sum of these pseudo-rewards with a function that depends on the quantity and the source of sign-ups under both $\texttt{OPT}$ and $\texttt{AC}$. By combining these lemmas, we establish a (non-linear) relationship between the expected value of $\texttt{AC}$ and that of $\texttt{OPT}$.

Lemma 3 (Lower Bound on $\texttt{AC}$ via Pseudo-rewards) *For any instance $\mathcal{I}$, the expected sum of all of the pseudo-rewards is a lower bound on the expected value of $\texttt{AC}$, i.e.,*

$$\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC}] \quad \geq \quad \mathbb{E}_{\boldsymbol{\omega}}\Big[\sum_{t\in[T]} L_t + \sum_{i\in[n]} K_i\Big], \tag{10}$$

where L_t and K_i are defined in (8) and (9), respectively.

[24] The set $\mathcal{V}^0$ is a function of the instance and the sample path, but we remind that we are suppressing that dependence.

The proof of Lemma 3 crucially relies on the fact that whenever internal traffic arrives, $\texttt{AC}$ recommends the opportunity which maximizes $\mu_{i,t}\psi(\mathrm{FR}_{i,t-1})$. Due to stochasticity in volunteers' realized sign-up decisions, this inequality holds only in expectation over all sample paths. We present the full proof in Appendix A.6.1.

In the subsequent lemma, we establish a lower bound on the expected sum of the pseudo-rewards. Recall that, for a fixed instance and sample path, we use counters such as $\texttt{AC}^{\text{INT}}_{i,T}$ to indicate the number of sign-ups for opportunity i made by volunteers $t \in \mathcal{V}^{\text{INT}}$ under the $\texttt{AC}$ algorithm. Similarly, we will use $\texttt{AC}^{0}_{i,T}$ to represent the amount of opportunity i's capacity filled by volunteers $t \in \mathcal{V}^0$ under the $\texttt{AC}$ algorithm. Mathematically, we have $\texttt{AC}^{0}_{i,T} = \sum_{t\in\mathcal{V}^0} \mathbb{1}[\tilde{\xi}_t(S^{\texttt{AC}}_t) = i]$. Furthermore, to mirror our notation for the $\texttt{AC}$ algorithm, we define $\texttt{OPT}^{\text{INT}}_{i,T}$ (resp. $\texttt{OPT}^{\text{EXT}}_{i,T}$) as the amount of opportunity i's capacity filled by internal traffic (resp. external traffic) under $\texttt{OPT}$ at the end of the horizon.

Lemma 4 (Lower Bound on Pseudo-Rewards) *For any instance $\mathcal{I}$, we have the following lower bound on the expected sum of all of the pseudo-rewards:*

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in[T]} L_t + \sum_{i\in[n]} K_i\right] \quad \geq \quad e^{-1/\underline{c}}\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} \texttt{AC}^{\text{EXT}}_{i,T} + \texttt{AC}^{0}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T}\cdot\psi\left(\frac{\texttt{AC}^{\text{INT}}_{i,T}}{c_i - \texttt{AC}^{\text{EXT}}_{i,T}}\right)\right.$$
$$\left. + c_i\left(1-\psi\left(\frac{\texttt{AC}^{\text{INT}}_{i,T} - \texttt{AC}^{0}_{i,T}}{c_i}\right) - 1/e\right)\right], \qquad (11)$$

where L_t and K_i are defined in (8) and (9), respectively.

Though we present (11) in expectation over all sample paths, in the proof of Lemma 4 we show that the inequality holds along each sample path by separately bounding the sum of the L_t pseudo-rewards and the sum of the K_i pseudo-rewards. The proof relies on properties of the function ψ, and the full proof details can be found in Appendix A.6.2.

Step 3: Bounding the Competitive Ratio of $\texttt{AC}$

The final step of the proof of Lemma 2 involves the creation of an instance-specific, factor-revealing mathematical program (MP) that serves as a lower bound on the ratio between $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC}]$ and $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{OPT}]$ on that instance. As we later elaborate upon, the program (MP) for instance $\mathcal{I}$ is designed such that we can construct a feasible solution using the outputs of $\texttt{AC}$ and $\texttt{OPT}$ on that instance.[25] The constraints are inspired by the lower bound on the expected value of $\texttt{AC}$ established in Step 2 as well as the physical constraints of the problem. The program (MP) partly consists of decision variables specific to each sample path $\boldsymbol{\omega}$ that can occur in instance $\mathcal{I}$. We use Ω to denote this set of sample paths, which has an associated probability measure induced by the instance $\mathcal{I}$.[26]

[25] We emphasize that (MP) depends on the instance $\mathcal{I}$, even though we suppress that dependence.

[26] We note that the probability measure is determined by a set of independent Bernoulli random variables.

We analyze (MP) via two additional lemmas. First, in Lemma 5, we show that the optimal value of (MP) is a lower bound on the ratio between the expected value of `AC` and the expected value of `OPT` in instance $\mathcal{I}$. To establish a lower bound on the competitive ratio of the `AC` algorithm, it then suffices to lower-bound the value of (MP) for all instances $\mathcal{I} \in \mathcal{I}_\beta$. To that end, in Lemma 6, we place a lower bound of z^* on the value of (MP), where we remind that z^* only depends on three properties of the instance $\mathcal{I}$: the EFET β, the minimum capacity $\underline{c}$, and the MCPR σ.

Table 1 Definition of the mathematical program (MP).

Given an instance $\mathcal{I}$, the inputs to (MP) are the set of opportunities $\mathcal{S}$, the EFET β, the MCPR σ, and the set of feasible sample paths Ω, along with its associated probability measure.

(MP) uses the set of variables $\vec{x} \in \mathbb{R}_{\geq 0}^{3\times n\times|\Omega|}$ and $\vec{y} \in \mathbb{R}_{\geq 0}^{2\times n\times|\Omega|} \setminus \vec{\mathbf{0}}$, along with $z \in [0,1]$

$$\min_{\vec{x},\vec{y},z} \quad z \qquad \textbf{(MP)}$$

$$\begin{aligned}
\text{s.t.} \quad & \forall i,\boldsymbol{\omega}, \quad c_i \geq y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}} \quad \textbf{(i)} \qquad c_i \geq x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}} \quad \textbf{(ii)} \qquad x_{2,i,\boldsymbol{\omega}} \geq x_{3,i,\boldsymbol{\omega}} \quad \textbf{(iii)} \\
& \qquad c_i = x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}} \quad \textbf{OR} \quad x_{1,i,\boldsymbol{\omega}} = y_{1,i,\boldsymbol{\omega}} \qquad \textbf{(iv)} \\
& \mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}}\right] \leq z \textstyle\sum_{i\in[n]} c_i \qquad \textbf{(v)} \\
& \mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}}\right] \geq (\beta - \sigma + z) \textstyle\sum_{i\in[n]} c_i \qquad \textbf{(vi)} \\
& \mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}} \cdot \psi\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - x_{1,i,\boldsymbol{\omega}}}\right) + c_i\left(1 - \psi\left(\frac{x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}}{c_i}\right) - 1/e\right)\right] \\
& \qquad \leq e^{1/\underline{c}} z \mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\right] \qquad \textbf{(vii)}
\end{aligned}$$

Lemma 5 (Lower-Bound on Ratio of Expected Values via (MP)) *For any instance $\mathcal{I}$, the ratio between the expected value of* `AC` *(i.e., $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC}]$) and the expected value of* `OPT` *(i.e., $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{OPT}]$) on instance $\mathcal{I}$ is at least the optimal value of* (MP)*.*

To prove Lemma 5, we consider the following candidate solution:

$$x_{1,i,\boldsymbol{\omega}} = \texttt{AC}_{i,T}^{\text{EXT}}, \qquad x_{2,i,\boldsymbol{\omega}} = \texttt{AC}_{i,T}^{\text{INT}}, \qquad x_{3,i,\boldsymbol{\omega}} = \texttt{AC}_{i,T}^{0},$$
$$y_{1,i,\boldsymbol{\omega}} = \texttt{OPT}_{i,T}^{\text{EXT}}, \qquad y_{2,i,\boldsymbol{\omega}} = \texttt{OPT}_{i,T}^{\text{INT}}, \qquad z = \frac{\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC}]}{\mathbb{E}_{\boldsymbol{\omega}}[\texttt{OPT}]}$$

Such a solution has an objective value equal to the ratio $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC}]/\mathbb{E}_{\boldsymbol{\omega}}[\texttt{OPT}]$ in (MP), and by construction it satisfies all constraints. We formally prove the feasibility of this solution in Appendix A.6.3. For intuition, note that the first two constraints in (MP) represent physical constraints of the problem – for any opportunity i, the amount of filled capacity cannot exceed c_i. The third constraint is satisfied according to the definition of our candidate solution and noting that $\mathcal{V}^0 \subseteq \mathcal{V}^{\text{INT}}$. The fourth constraint encodes the following property of `AC`: for any opportunity i, `AC` fills less capacity with external traffic compared to $\texttt{OPT}$ only if opportunity i has already reached capacity under $\texttt{AC}$. The fifth constraint holds by the definition of z and the fact that $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{OPT}] \leq \sum_{i\in[n]} c_i$. The sixth constraint establishes a lower bound on the capacity filled by volunteers in $\mathcal{V}^{\text{EXT}}$ and $\mathcal{V}^0$, which we remind are the two sets of volunteers for which $\texttt{OPT}$ could not have made a better decision than $\texttt{AC}$ (see (8) and the following discussion). Establishing this lower bound is technically involved, and we defer the details to Appendix A.6.3. For now, we simply highlight that the bound depends on the EFET β, the MCPR σ, and the ratio between the expected values of $\texttt{AC}$ and $\texttt{OPT}$. The seventh and final constraint is satisfied as a direct consequence of Lemmas 3 and 4 in Step 2.

In general, (MP) is non-convex. Despite this, in the following lemma, we are able to bound the optimal value of (MP) for any instance.

Lemma 6 (Lower Bound on the Optimal Value of (MP)) *For any instance $\mathcal{I}$, the optimal value of* (MP) *is at least z^*, where z^* is defined in* (6) *in the statement of Theorem 2.*

The proof of Lemma 6 is mainly algebraic and relies on repeatedly relaxing the program's constraints and restricting its domain until we can ultimately establish a lower bound of z^*. We defer the details to Appendix A.6.4. In combination, the two lemmas of Step 3 establish a lower bound on the ratio $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC}]/\mathbb{E}_{\boldsymbol{\omega}}[\texttt{OPT}]$ for any instance $\mathcal{I} \in \mathcal{I}_\beta$, where the bound depends on only the EFET β, the minimum capacity $\underline{c}$, and the MCPR σ.

Together, these three steps prove Lemma 2, namely, that z^* is a lower bound on the competitive ratio of the $\texttt{AC}$ algorithm. In combination with Lemma 1, we have shown that the competitive ratio of the $\texttt{AC}$ algorithm is at least $f(\beta, \underline{c}, \sigma)$, as defined in the statement of Theorem 2. □

5. Model Extensions

In many practical settings, platforms can provide more than one recommendation to internal traffic, often in the form of a ranking. Here, we discuss the ways in which our model and results can generalize to such settings, which we henceforth refer to as the *ranking setting.*

We begin this section by describing how we augment the model of Section 3. Upon the arrival of an internal traffic volunteer, we now allow the platform to present a ranking of opportunities $\vec{S} \in \mathcal{S}^{\mathcal{R}}$, instead of a single recommendation.[27] The volunteer *views* (at most) one opportunity from this ranked subset.[28] As before, the volunteer will *sign up* for the viewed opportunity with their pair-specific conversion probability. We use $\phi_{i,t}(\vec{S})$ to denote the probability that volunteer t signs up for opportunity i when presented with the ranking $\vec{S}$. (We augment each volunteer's type to

[27] We allow the domain of possible rankings $\mathcal{S}^{\mathcal{R}}$ to consist of arbitrary ranked subsets of opportunities. We only require that it includes the singleton $\{0\}$, which deterministically results in no sign-up from that volunteer.

[28] For external traffic, we continue to follow the convention that any algorithm must recommend the single targeted opportunity i_t^*, which is then directly viewed by the volunteer.

include any parameters necessary to fully specify these probabilities for every possible ranking.) We use the random variable $\xi_t(\vec{S})$ to denote the volunteer's sign-up decision when presented with the ranking $\vec{S}$, which is either 0 or the opportunity viewed by the volunteer.

Our benchmark `OPT` (see Definition 1) generalizes to the ranking setting by simply recommending the optimal ranked subset of opportunities to arriving internal traffic, which can again be found by solving a dynamic program of exponential size.[29] Likewise, the `AC` algorithm naturally generalizes to an algorithm that we denote by `AC-R`. The `AC-R` algorithm follows exactly the same steps as the `AC` algorithm (see Algorithm 2), except instead of recommending the single opportunity i that maximizes $\mu_{i,t}\psi(\mathrm{FR}_{i,t-1})$, the `AC-R` algorithm recommends the ranking $\vec{S}_t^{\texttt{AC-R}}$ that satisfies

$$\vec{S}_t^{\texttt{AC-R}} \in \operatorname{argmax}_{\vec{S}\in\mathcal{S}^{\mathcal{R}}} \sum_{i\in[n]} \phi_{i,t}(\vec{S})\cdot\psi(\mathrm{FR}_{i,t-1}). \tag{12}$$

Henceforth, we assume that the platform can efficiently solve (12), which is a common assumption in the literature (Golrezaei et al. 2014, Gong et al. 2021). Given this assumption, we are able to establish results that are similar to Theorem 2, as formalized in the following proposition.

Proposition 4 (Lower Bound on the Competitive Ratio of `AC-R`) *Let the smallest capacity be given by $\underline{c}$. Then, for any effective fraction of external traffic β, the competitive ratio of* `AC-R` *algorithm is at least* $\max\{\beta, e^{-1/\underline{c}}(1-1/e)\}$.

This lower bound on the competitive ratio of the `AC-R` algorithm is numerically equivalent to the lower bound established in Theorem 2 when the MCPR σ exceeds $e-1$ (beyond which the lower bound is constant in σ). The intuition developed in Section 4.3 applies in this setting, too: we cannot guarantee that the `AC-R` algorithm fills *any* capacity with external traffic unless the EFET is sufficiently large. In fact, the instance of Example 1 is a special case of the ranking setting (where the platform recommends one opportunity which the volunteer deterministically views). Thus, the lower bound of Proposition 4 cannot be improved, at least for that set of parameters ($\beta = 1-1/e,\ \underline{c}\to\infty,\ \sigma\to\infty$). Furthermore, in the ranking setting, we cannot necessarily improve our result even when the MCPR is bounded. In our base model, the MCPR σ bounds the "relative value" of two different (non-empty) recommendations (i.e., the ratio of their expected number of sign-ups). However, in the ranking setting, the "relative value" of two different recommendations can be quite large, regardless of the MCPR. We defer the proof of Proposition 4 to Appendix B.1.

Though the result of Proposition 4 holds for arbitrary choice functions, our proof technique is flexible enough to (potentially) provide stronger results when tailored to a particular choice

[29] For any algorithm with an optimality criteria (such as `AC` and `OPT`), in the presence of multiple optimal solutions, we follow the convention of choosing the optimal solution that presents the ranked subset of the smallest size, breaking ties in favor of the solution that lexicographically minimizes the indices of the ranked subset.

function. For example, consider a special case of the *cascade* (or sequential search) model for volunteer choice, which has been used to model search on online platforms (see, e.g., Aggarwal et al. 2008 and Kempe and Mahdian 2008).

Definition 5 (Opportunity-Agnostic Cascade Model) *The opportunity-agnostic cascade model is parameterized by a volunteer-specific view probability $\nu_t > 0$ and a volunteer-specific exit probability $q_t \geq 0$. Given a ranked subset of opportunities (of length at most K), the volunteer sequentially "examines" the opportunities starting from the top (i.e., position 1). The volunteer views the top-ranked opportunity independently with probability ν_t. Conditional on not viewing the opportunity, the volunteer exits the platform independently with probability q_t. If the volunteer does not exit, they repeat the same process for the second-ranked opportunity, and so on. If the volunteer reaches the end of the ranked list without viewing an opportunity, they exit the platform.*

The opportunity-agnostic cascade model is a special case of the cascade model in which the view probabilities depend only on the ranked position of an opportunity, and are "agnostic" to the identity of the opportunity itself. This property leads to the following observation: under the opportunity-agnostic cascade model, ranking opportunities in descending order of $\mu_{i,t} \cdot \psi(\mathrm{FR}_{i,t-1})$ satisfies `AC-R`'s optimality condition (as given in (12)). Using this critical observation (formalized in Claim 8 of Appendix B.2), we are able to strengthen Proposition 4 under this choice model.

Proposition 5 (`AC-R` Under the Opportunity-Agnostic Cascade Model) *Let the smallest capacity be given by $\underline{c}$, let the maximum conversion probability ratio (given by Definition 4) be at most σ, and suppose each volunteer choice follows the opportunity-agnostic cascade model (specified in Definition 5). Then, for any effective fraction of external traffic β, the competitive ratio of the* `AC-R` *algorithm is at least $f(\beta, \underline{c}, \sigma)$, as defined in the statement of Theorem 2.*

The proof of Proposition 5 (deferred to Appendix B.2) crucially relies on the fact that the probability of viewing an opportunity depends only on its position in the ranking. Therefore, different rankings can only have different "relative values" if either (a) there are differences in conversion probabilities *conditional* on a view, or (b) the rankings are of different length. The former influences our bound via the MCPR σ, while we account for the latter by leveraging the observation that the `AC-R` algorithm ranks opportunities in descending order of $\mu_{i,t} \cdot \psi(\mathrm{FR}_{i,t-1})$.

6. Evaluating Algorithm Performance on VM Data

In this section, we use data from VM to numerically evaluate the performance of the `AC` algorithm in instances closer to practice. Section 6.1 provides useful background on the VM platform. Section 6.2 explains how we use VM data to construct instances of our model. Finally, Section 6.3 compares the performance of `AC` to various benchmarks and demonstrate the effectiveness of our algorithm.

6.1. VolunteerMatch Background

As mentioned in Section 1, VM is the U.S.'s largest online platform for connecting volunteers and opportunities. To carry out our case study, we draw upon VM's database – which provides us with information on opportunities' and volunteers' activities, such as sign-ups– as well as VM's Google Analytics (GA) dataset, which consists of a subset of session-level website traffic activity.

Organizations looking for volunteers post an opportunity, and each opportunity has a location (in-person or virtual), timing (specific dates/times or a flexible schedule), number of volunteers needed, and up to three associated "causes," out of a list of thirty (e.g., LGBTQ, seniors, hunger, etc.). In Figure 3, we display the percentage of opportunities that are associated with each cause. For the purpose of this section, we will only consider the 10,737 virtual opportunities appearing in our GA data between August 2020 and March 2021 for which we have precise data on capacity. (We provide more details about this subset of opportunities in Appendix C.1.)

Volunteers also select a subset of the different causes when creating an account on VM, and Figure 3 displays the percentage of volunteers interested in each cause. When internal traffic visits the VM website, they are presented with a ranked list of opportunities, which can be filtered to include only opportunities associated with the causes they are interested in. Volunteers then can *view* an opportunity, i.e, they can click on an opportunity and learn more about the job description, requirements, etc. Nearly half (45%) of internal traffic leaves the site without viewing any opportunity. By contrast, all external traffic volunteers go directly to view their targeted opportunity. Conditional on viewing an opportunity, a volunteer may choose to sign up for it.

As discussed in Section 1, opportunities differ greatly in the amount of sign-ups from both internal and external traffic that they currently receive, which results in some opportunities receiving excessive sign-ups. Our numerical study aims to understand whether using the `AC` algorithm on the VM platform would reuslt in a better utilization of traffic. To that end, in the following section, we construct an instance based on VM data, and we test `AC` on this instance against various benchmarks, including a proxy for current practice on the site. We then investigate how the performance of `AC` changes as a function of the EFET by considering a family of similar instances (which loosely represent settings where VM has some control over the destination of external traffic).

6.2. Instance Construction

We now briefly describe how we use the available data to construct an instance that is as close to practice as possible. More details on each component of the instance can be found in Appendix C.3.

Set of Opportunities. Because we wish to compare the performance of `AC` against that of an offline benchmark (which is a computationally-intensive task, as described below and in Appendix C.4), we focus our numerical study on a random sample of 100 opportunities from the set of 10,737 opportunities described in Section 6.1 and include relevant robustness checks in Appendix C.2. We have precise data on the number of volunteers needed for each of these opportunities, which we use as the opportunities' capacities.

Source and Arrival Order. Based on volunteer activity, we estimate that there are $T = 11,345$ website visitors for our subset of 100 opportunities,[30] 83% of whom are internal traffic (recall that 45% of these volunteers leave the site without viewing anything). To generate the arrival sequence, we preserve the traffic pattern observed in the session-level data (after appropriate scaling, as we only observe session-level data for approximately 20% of website traffic). For example, if volunteer 1 is external traffic and volunteer 2 is internal traffic in the dataset, then in our simulation, the first five volunteers will all be external traffic, and the next five will be internal traffic.

Volunteer Conversion Probabilities. External traffic volunteers are straightforward: each such volunteer goes directly to view their targeted opportunity and may choose to sign-up for this opportunity; we directly observe these sign-up realizations in the data and preserve these realizations in our simulation (after appropriate scaling as described above).

By contrast, internal traffic volunteers could view any one of a ranked list of opportunities. We assume that internal traffic volunteers only have non-zero pair-specific conversion probabilities for opportunities for which they are *compatible*; however, we do not directly observe this compatibility from the data. To estimate this compatibility structure, we use data on volunteer and opportunity causes.[31] As shown in Figure 3, there is significant variation in volunteers' interest across causes as well as variation in the number of opportunities associated with each cause. To construct the compatibility for each internal traffic volunteer, we randomly generate causes for these volunteers proportional to the empirical distribution we observe, independently across causes and volunteers. In addition, we preserve the causes associated with each opportunity in our sample. A volunteer is compatible with an opportunity if and only if they share at least one cause in common.

To estimate the vector of conversion probabilities for internal traffic volunteers conditional on viewing a compatible opportunity, we run a logistic regression on the observed view conversion probabilities using opportunity-level characteristics (e.g., causes). For all incompatible volunteer and opportunity pairs, we set the conversion probability to zero. Below we discuss how internal traffic volunteers make viewing decisions among compatible opportunities.

---

[30] There are significantly more visitors to the VM site over the 8-month period that we study; this represents the number of visitors proportional to this subset of 100 random opportunities. See Appendix C.3 for more details.

[31] For in-person opportunities, compatibility also depends on location, hence our focus on virtual opportunities.

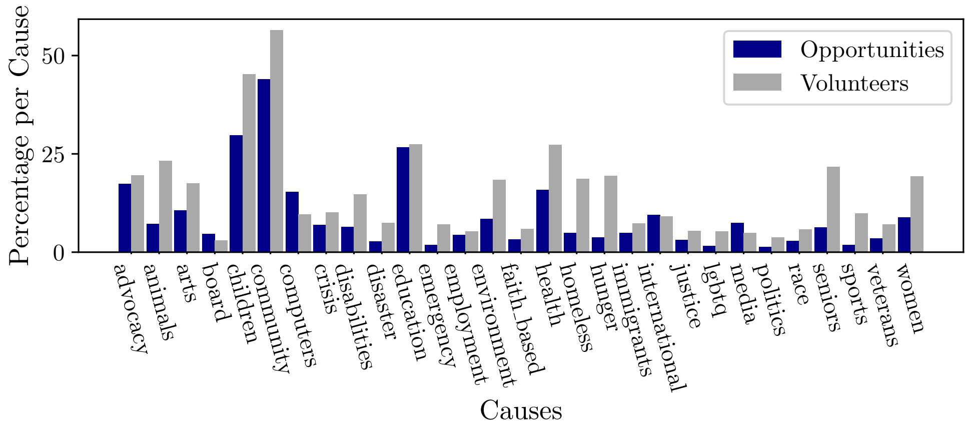


Figure 3 **Percentage of volunteers and opportunities associated with each cause. These percentages need not sum to 100%, as opportunities and volunteers can be associated with multiple causes.**

Volunteer Choices. Internal traffic volunteers arriving to VM are presented with a ranked list of opportunities. They can choose to view an opportunity on the list and, after viewing, they may choose to sign up for it. While our dataset is rich enough to allow us to construct all of the aforementioned dimensions of our instance, it does not contain precise information on the ranking presented to each internal traffic volunteer. Consequently, we cannot infer the choice behavior of these volunteers through the available VM data. Thus, in this simulation study, we assume internal traffic behaves according to the opportunity-agnostic cascade model introduced in Definition 5.

Specifically, we assume volunteers only consider the top three options presented (i.e., $K = 3$). Before exhausting those three options, we assume the view and exit probabilities are homogeneous and equal to $\nu = 0.3$ and $q = 0.24$, respectively. These parameters are chosen to match our empirical observation that only 55% of internal traffic views an opportunity. Upon viewing an opportunity, the volunteer will sign up according to their conversion probability, and will otherwise depart.

Benchmarks. To gain a better understanding of the performance of `AC`, we compare it with the following alternatives:

Current Practice (`CP`): Under our first benchmark – which serves as a stylized proxy for VM's current practice – the platform ranks compatible opportunities based on the "recency" of an opportunities's actions.[32] Thus, it does not account for (i) opportunities' current fill rates and (ii) the traffic source (i.e., internal or external).

Smart Current Practice (`SCP`): Our second benchmark is a slightly more sophisticated version of `CP` that only ranks opportunities that are not yet full. That is, this algorithm ranks compatible opportunities with remaining capacity sorted by recency.

[32] Once an organization posts an opportunity on VM's platform, it can take various actions to modify the opportunity, affecting its recency. We preserve each opportunity's actions from VM's internal dataset, and we then recreate the ranking that internal traffic would see under this algorithm, which we remind is a proxy for current practice.

Algorithm 1 (`MSVV`): As a third benchmark, we use the `MSVV` algorithm introduced in Mehta et al. (2007). As discussed before, this algorithm makes ranking decisions based on conversion probabilities and the current fill rates.

Upper bound on `OPT` ($\overline{\texttt{OPT}}$): Our benchmark `OPT` (see Definition 1) is a dynamic program of exponential size and is therefore challenging to compute. As is standard in the literature (see, e.g., Alaei et al. 2013), we upper-bound `OPT` with the solution to a deterministic fractional matching, denoted $\overline{\texttt{OPT}}$, formally defined in Appendix C.4. We use $\overline{\texttt{OPT}}$ in lieu of `OPT` as a normalization factor to evaluate the performance of `AC` and the other benchmarks.

6.3. Results

We now aim to understand how well `AC` performs on the data-driven instances constructed as described in Section 6.2. Specifically, we first examine the performance of `AC` against the aforementioned benchmarks. Then, we investigate how the performance of `AC` changes under different levels of the EFET, which a platform may be able to achieve if it has some control over external traffic.

Performance of `AC`: In Figure 4a, we present the value of `AC` and the three benchmarks introduced above (`CP`, `SCP`, and `MSVV`), averaged over 50 simulations[33] and normalized by $\overline{\texttt{OPT}}$. Figure 4a shows that `AC` dramatically outperforms `CP` and achieves 85% of $\overline{\texttt{OPT}}$. Given that $\overline{\texttt{OPT}}$ itself is an upper-bound on `OPT`, this implies that the performance of `AC` is remarkably close to that of `OPT`. Though `AC` also outperforms `SCP`, accounting for fill rates even in a binary way (i.e., by not showing opportunities that have reached capacity) significantly narrows the gap.[34] Finally, `AC` performs similarly to `MSVV`. We highlight that only 19% of the capacity of these opportunities could be filled by external traffic in this instance, i.e., EFET= 19%. Hence, this is consistent with our theoretical results, as there is no gap between the competitive ratios of `AC` and `MSVV` when the EFET is low.[35]

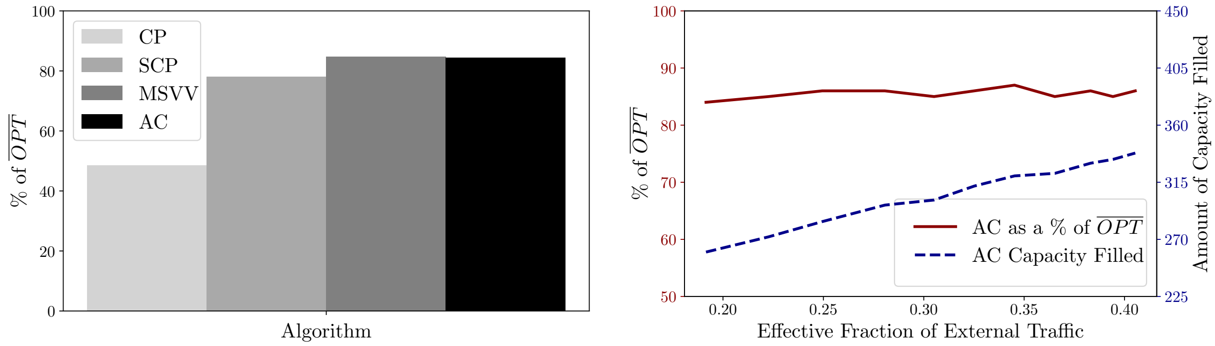
`AC`'s strong performance is due to its effectiveness in re-distributing internal traffic. We illustrate this in Figure 5, which shows the sign-up distributions resulting from a single simulated run of `CP` and `AC`. We normalize each opportunity's sign-ups by its capacity such that the black line represents its sufficient number of sign-ups. `CP` results in a highly non-uniform sign-up distribution as it fails to "de-prioritize" those that have already received a sufficient amount of sign-ups. Under `CP`, we see that in hindsight 83% of the internal traffic can be re-distributed. In comparison, under `AC` the sign-ups from internal traffic are distributed more evenly across opportunities. Consequently, the amount of internal traffic that can be re-distributed in hindsight decreases from 83% to 18%.

---

[33] In each simulation, whenever an internal traffic volunteer arrives, we draw random variables to determine which opportunity a volunteer views and (conditional on viewing an opportunity) whether the volunteer signs up.

[34] In Appendix C.2 we show how these performance gaps scale as a function of the instance size.

[35] However, unlike in our theoretical results, the arrival pattern in this instance is not adversarially chosen.



(a) (b)

Figure 4 **(a) Performance of `AC` and three benchmarks `CP`, `SCP`, and `MSVV` (b) performance of `AC` under different levels of the EFET.**

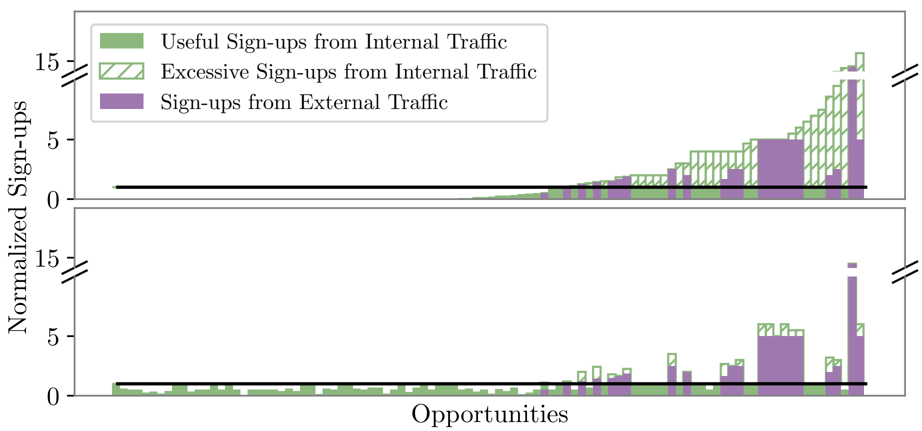


Figure 5 **Distribution of normalized sign-ups across a subset of opportunities under `CP` (top) and `AC` (bottom).**

Varying EFET: We conclude this case study with a thought experiment that can be useful to inform the design of online platforms such as VM, and which allows us to investigate the performance of `AC` as the EFET changes. As discussed in Section 1, external traffic can be the result of targeted outreach activities by organizations, but it can also be driven by the platform. For example, VM sends out recurring emails to its subscribers to highlight a few opportunities.

As an illustration of the potential benefit if the platform had some control over the destination of external traffic, consider the top panel of Figure 5. Even without internal traffic, 65% of the sign-ups from external traffic are excessive. (We remind that this figure illustrates the output of a single simulated run of `CP` and `AC` in the instance constructed in Section 6.2). With this in mind, we consider the following family of perturbed instances. We define a parameter $\eta \in (0, 1)$; for each opportunity that has *excess* external traffic in hindsight (i.e., more sign-ups from external traffic than its total capacity), we take an η fraction of that excess external traffic and we re-assign it to opportunities selected uniformly at random from among those that do not have excess external traffic in hindsight and that have at least one cause in common with that opportunity. For example, suppose opportunity a received 10 external sign-ups, but has capacity 5. This opportunity has 5 excess external sign-ups. If $\eta = 0.2$, we would take one of those sign-ups and allocate it to a different opportunity. As we vary η from 0 to 1, the EFET increases from its original value of 19% to 41%.

Figure 4b shows the expected value of `AC` (in blue) and its expected value normalized by $\overline{\texttt{OPT}}$ (in red) as a function of the EFET. As η (or equivalently, the EFET) increases, the amount of filled capacity under `AC` substantially increases. However, $\overline{\texttt{OPT}}$ also increases, such that their ratio remains almost unchanged. The fact that the normalized performance of `AC` is largely unchanged for an EFET between 19% and 41% is consistent with our theoretical results, as the competitive ratio only improves when the EFET exceeds $1/e$ (see Theorem 2 and Figure 2a). In practice, of course, re-directing external traffic is more nuanced than simply re-allocating an external sign-up uniformly at random, as we do in this simulation. We use this approach as a crude way to simultaneously investigate the performance of `AC` as the EFET changes and to illustrate the potential benefit of designing off-platform outreach (such as targeted emails). Studying how to optimally design external traffic in a well-motivated setting is an interesting potential direction for future research.

7. Conclusion

In this paper, we introduce a framework for making online recommendations to maximize matches in the presence of external traffic, motivated by platforms such as VolunteerMatch (the largest online volunteer engagement network in the US, and our industry partner). Our recommendation algorithm, Adaptive Capacity (`AC`), does not know the amount of external traffic *a priori*, yet it nevertheless provides strong parameterized guarantees (relative to both the commonly-used `MSVV` algorithm and the upper bound we establish on any online algorithm). Our flexible analysis allows us to generalize our results to settings where the platform provides a ranked set of recommendations. Beyond theoretical guarantees, we demonstrate `AC`'s practical effectiveness in simulations based on VM data. We are currently collaborating with VM to implement a version of our algorithm.

More generally, our work shows the importance of accounting for the source of traffic in decision-making on platforms with multi-channel traffic, which opens up opportunities for further research. For instance, while we have focused on settings where the platform cannot influence external traffic, some platforms may have some degree of control over the timing or the destination of this traffic (e.g., via marketing campaigns or curated email recommendations). Also, platforms with external traffic may have objectives beyond maximizing the number of matches (e.g., platforms such as DonorsChoose may aim to maximize the number of donation campaigns that reach a certain threshold). Studying the platform design in such settings is an interesting direction for future work.

References

Gagan Aggarwal, Jon Feldman, Shanmugavelayutham Muthukrishnan, and Martin Pál. Sponsored search auctions with markovian users. In *International Workshop on Internet and Network Economics*, pages 621–628. Springer, 2008.

Saeed Alaei, Mohammad Taghi Hajiaghayi, and Vahid Liaghat. The online stochastic generalized assignment problem. In *APPROX-RANDOM*, 2013.

Saeed Alaei, Ali Makhdoumi, Azarakhsh Malekian, and Saša Pekeč. Revenue-sharing allocation strategies for two-sided media platforms: Pro-rata vs. user-centric. *Management Science*, 2022.

Ali Aouad and Daniela Saban. Online assortment optimization for two-sided matching platforms. *Available at SSRN 3712553*, 2020.

Nick Arnosti and Peng Shi. Design of lotteries and wait-lists for affordable housing allocation. *Management Science*, 66(6):2291–2307, 2020.

Itai Ashlagi, Anilesh K Krishnaswamy, Rahul Makhijani, Daniela Saban, and Kirankumar Shiragur. Assortment planning for two-sided sequential matching markets. *arXiv preprint arXiv:1907.04485*, 2019.

Santiago Balseiro, Haihao Lu, and Vahab Mirrokni. The best of many worlds: Dual mirror descent for online allocation problems. *arXiv preprint arXiv:2011.10124*, 2020.

Omar Besbes, Francisco Castro, and Ilan Lobel. Surge pricing and its spatial supply response. *Management Science*, 67(3):1350–1367, 2021.

Niv Buchbinder, Kamal Jain, and Joseph Seffi Naor. Online primal-dual algorithms for maximizing ad-auctions revenue. In *European Symposium on Algorithms*, pages 253–264. Springer, 2007.

Niv Buchbinder, Joseph Seffi Naor, et al. The design of competitive online algorithms via a primal–dual approach. *Foundations and Trends® in Theoretical Computer Science*, 3(2–3):93–263, 2009.

Antoine Désir, Vineet Goyal, and Jiawei Zhang. Capacitated assortment optimization: Hardness and approximation. *Operations Research*, 2021.

Daria Dzyabura and Srikanth Jagabathula. Offline assortment optimization in the presence of an online channel. *Management Science*, 64(6):2767–2786, 2018.

Adam N Elmachtoub, Vineet Goyal, Roger Lederman, and Harsh Sheth. Revenue management with product retirement and customer selection. *Available at SSRN 4033922*, 2022.

Hossein Esfandiari, Nitish Korula, and Vahab Mirrokni. Online allocation with traffic spikes: Mixing adversarial and stochastic models. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 169–186, 2015.

Jacob Feldman and Danny Segev. The multinomial logit model with sequential offerings: Algorithmic frameworks for product recommendation displays. *Operations Research*, 2022.

Yiding Feng, Rad Niazadeh, and Amin Saberi. Linear programming based online policies for real-time assortment of reusable resources. 2019.

Negin Golrezaei, Hamid Nazerzadeh, and Paat Rusmevichientong. Real-time optimization of personalized assortments. *Management Science*, 60(6):1532–1551, 2014.

Xiao-Yue Gong, Vineet Goyal, Garud N Iyengar, David Simchi-Levi, Rajan Udwani, and Shuangyu Wang. Online assortment optimization with reusable resources. *Management Science*, 2021.

Vineet Goyal and Rajan Udwani. Online matching with stochastic rewards: Optimal competitive ratio via path based formulation. *arXiv preprint arXiv:1905.12778*, 2019.

Vineet Goyal, Garud Iyengar, and Rajan Udwani. Asymptotically optimal competitive ratio for online allocation of reusable resources. *arXiv preprint arXiv:2002.02430*, 2020.

Dawsen Hwang, Patrick Jaillet, and Vahideh Manshadi. Online resource allocation under partially predictable demand. *Operations Research*, 2021.

Nicole Immorlica, Brendan Lucier, Vahideh Manshadi, and Alex Wei. Designing approximately optimal search on matching platforms. *Available at SSRN 3850164*, 2021.

Bala Kalyanasundaram and Kirk R Pruhs. An optimal deterministic algorithm for online b-matching. *Theoretical Computer Science*, 233(1-2):319–325, 2000.

Yash Kanoria and Daniela Saban. Facilitating the search for partners on matching platforms. *Management Science*, 2021.

Richard M Karp, Umesh V Vazirani, and Vijay V Vazirani. An optimal algorithm for on-line bipartite matching. In *Proceedings of the twenty-second annual ACM symposium on Theory of computing*, pages 352–358, 1990.

David Kempe and Mohammad Mahdian. A cascade model for externalities in sponsored search. *Internet and Network Economics*, pages 585–596, 2008.

Will Ma and David Simchi-Levi. Algorithms for online matching, assortment, and pricing with tight weight-dependent competitive ratios. *Operations Research*, 68(6):1787–1803, 2020.

Vahideh Manshadi and Scott Rodilitz. Online policies for efficient volunteer crowdsourcing. *Management Science*, 2022.

Aranyak Mehta and Debmalya Panigrahi. Online matching with stochastic rewards. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 728–737. IEEE, 2012.

Aranyak Mehta, Amin Saberi, Umesh Vazirani, and Vijay Vazirani. Adwords and generalized online matching. *Journal of the ACM (JACM)*, 54(5):22–es, 2007.

Aranyak Mehta et al. Online matching and ad allocation. *Foundations and Trends® in Theoretical Computer Science*, 8(4):265–368, 2013.

Joseph Naor and David Wajc. Near-optimum online ad allocation for targeted advertising. *ACM Transactions on Economics and Computation (TEAC)*, 6(3-4):1–20, 2018.

Ignacio Ríos, Daniela Saban, and Fanyin Zheng. Improving match rates in dating markets through assortment optimization. *Available at SSRN*, 2020.

Paat Rusmevichientong, Mika Sumida, and Huseyin Topaloglu. Dynamic assortment optimization for reusable products with random usage durations. *Management Science*, 2020.

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

Rajan Udwani. Adwords with unknown budgets. *arXiv preprint arXiv:2110.00504*, 2021.

Andrew Chi-Chin Yao. Probabilistic computations: Toward a unified measure of complexity. In *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)*, pages 222–227. IEEE Computer Society, 1977.

Appendix A: Omitted Proofs of Section 4

A.1. Proof of Proposition 1 (Section 4.1)

The proof of Proposition 1 follows from the more general hardness result of Theorem 1, which establishes an upper bound of $1 - 1/e$ on the competitive ratio of any online algorithm in the special case where there is no external traffic(i.e., when $\beta = 0$).

We start from the instance that establishes this result ($\mathcal{I}_2(0)$, described in Appendix A.4), which consists of a total capacity of NC and an equal number of internal traffic volunteers. Fixing a particular $\beta \in [0, 1)$,[36] we add one opportunity to that instance with capacity $\frac{\beta}{1-\beta}NC$. To exactly fill this opportunity, we append $\frac{\beta}{1-\beta}NC$ external traffic volunteers to the start of the arrival sequence, where each of these arriving volunteers has a conversion probability of 1 for the newly-added opportunity.

By design, (i) all external traffic arrives first, (ii) the EFET is exactly equal to β, (iii) the new opportunity will be entirely filled with external traffic under any algorithm, as this traffic directly views the opportunity, but (iv) by Theorem 1, no online algorithm can achieve a competitive ratio better than $1 - 1/e$ on the remaining opportunities (none of the added volunteers are compatible with the remaining opportunities). Putting these four observations together, we have established an upper bound of $\beta + (1-\beta)(1-1/e)$ on the competitive ratio of any online algorithm when the external traffic arrives first.[37]

A.2. Proof of Proposition 2 (Section 4.1)

Consider a family of instances $\mathcal{I}_1(\beta)$ parameterized by the EFET β. In each instance, there are a large number of opportunities N, each with identical large capacity C. The arrival sequence consists of NC volunteers, and for a given effective fraction of external traffic β, the first βNC of these volunteers are external traffic.[38] All volunteers have conversion probabilities of 1 or 0, and if $\mu_{i,t} = 1$ (resp. 0), we will refer to opportunity i and volunteer t as *compatible* (resp. incompatible).

To help describe the compatibility structure of the arriving volunteers, we first define constants $\hat{\alpha}_1$ and $\hat{\alpha}_2$, where the former is the unique solution in $[0, 1]$[39] to

$$\beta = \hat{\alpha}_1 + (1 - \hat{\alpha}_1)\Big(\exp\big(-\hat{\alpha}_1/(1-\hat{\alpha}_1)\big) - 1\Big),$$

and the latter is defined as

$$\hat{\alpha}_2 = 1 - \frac{1 - \hat{\alpha}_1}{\exp\left(\exp(-\hat{\alpha}_1/(1-\hat{\alpha}_1))\right)}.$$

We illustrate the arrival sequence (and its associated compatibility structure) for this family of instances in Figure 6. To be precise, the βNC external traffic volunteers arrive first, and for each opportunity $i \in$

[36] For $\beta = 1$, we have the trivial result that the upper bound on the competitive ratio is 1.

[37] To show that this upper bound holds for any minimum capacity $\underline{c}$, it suffices to add an additional opportunity with capacity $\underline{c}$ for which volunteers have conversion probability of 0. The value of `OPT` and the upper bound on the performance of any algorithm do not change, and the EFET also remains the same in the limit as N approaches infinity.

[38] We assume that $(1-\beta)NC$ is an integer. This assumption does not impact the upper bound in the statement of Proposition 2, as the expression comes from taking the limit as N approaches ∞.

[39] We note that for any $\beta \in [0, 1]$, it is easy to verify algebraically that there is a unique solution in the interval $[0, 1]$ for $\hat{\alpha}_1$.

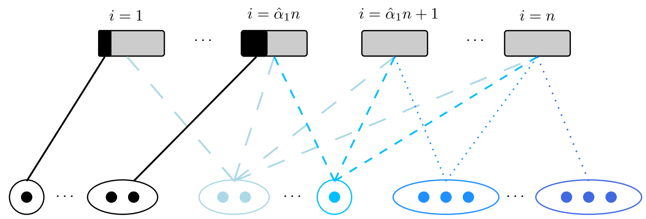


Figure 6 **The family of instances generating the upper bound on MSVV when all external traffic arrives first.**

$\{1,\ldots,\hat{\alpha}_1 N\}$, there are $C\left(1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^i\right)$ compatible external traffic arrivals for that opportunity. After the arrival of the last external traffic, the internal traffic arrives, according to the following compatibility structure: for each opportunity $i\in[N]$, there is a batch of Δ_i sequentially-arriving homogeneous volunteers. For each $i\in\{1,\ldots,\hat{\alpha}_1 N\}$, there are $\Delta_i = C\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^i$ volunteers who are compatible with all opportunities $j\geq i$. In addition, for each $i\in\{\hat{\alpha}_1 N+1,\ldots,N\}$, there are $\Delta_i = C$ volunteers who are again compatible with all opportunities $j\geq i$.

First, we verify that the EFET is equal to β in the limit as N gets large.

$$\begin{aligned}
\frac{1}{NC}\sum_{i=1}^{\hat{\alpha}_1 N} C\left(1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^i\right) &= \frac{1}{NC}\sum_{i=1}^{\hat{\alpha}_1 N}\left[C\left(1-\left(1-\frac{1}{(1-\hat{\alpha}_1)N+1}\right)^i\right)\right] \\
&\xrightarrow{N\to\infty} \int_0^{\hat{\alpha}_1}\left[1-\exp\left(\frac{-x}{1-\hat{\alpha}_1}\right)\,\partial x\right] \qquad (13)\\
&= \left(\hat{\alpha}_1 + (1-\hat{\alpha}_1)\left(\exp\left(\frac{-\hat{\alpha}_1}{1-\hat{\alpha}_1}\right)-1\right)\right) \\
&= \beta \qquad (14)
\end{aligned}$$

In (13), we use the fact that $(1-1/n)^{nx}$ approaches $\exp^{-x}$ as n approaches infinity. Furthermore, (14) follows by applying the definition of $\hat{\alpha}_1$. Next, we analyze the value of MSVV and OPT on the above family of instances via the following two claims.

Claim 1 *For any EFET β, the fraction of total capacity filled under* MSVV *on $\mathcal{I}_1(\beta)$ is at most $\hat{\alpha}_2$.*

Proof of Claim 1 To prove this claim, we will bound the amount of filled capacity for each opportunity under MSVV. First, we will show that the $\hat{\alpha}_1 N$ opportunities that receive external traffic do not receive any matches from internal traffic; i.e., for each $i\in[\hat{\alpha}_1 N]$, we will show that $\texttt{MSVV}_{i,T} = C\left(1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^i\right)$. Suppose towards a contradiction that there exists some opportunity $j\in[\hat{\alpha}_1 N]$ which receives a match from internal traffic under MSVV. Due to restrictions on compatibility, this match must have come from one of the first j batches of internal traffic, which in total represents

$$\sum_{i=1}^{j}\Delta_i = \sum_{i=1}^{j} C\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^i = C\left((1-\hat{\alpha}_1)N\right)\left(1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^j\right) \qquad (15)$$

internal traffic volunteers. We are supposing that one of these volunteers was allocated to opportunity j. In that case, due to the pigeonhole principle, there must be at least one opportunity j' – from among the $(1-\hat{\alpha}_1)N$ opportunities that *did not* receive external traffic – with a filled capacity strictly less than $C\left(1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^j\right)$ upon the arrival of the last volunteer in batch j. By definition, MSVV should never have recommended j ahead of j', giving us a contradiction.

Next, we show that each opportunity $i \in \{\hat{\alpha}_1 N+1,\ldots,N\}$ has a filled capacity of

$$\texttt{MSVV}_{i,T} = \min\left\{C, C\left(1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^{\hat{\alpha}_1 N} + \sum_{j=\hat{\alpha}_1 N+1}^{i} \frac{1}{N-j+1}\right)\right\}.$$

Note that in this matching setting, MSVV recommends opportunities to equalize their fill rate. Thus, after the arrival of the $\hat{\alpha}_1 N^{\text{th}}$ batch of volunteers, all opportunities $j \in \{\hat{\alpha}_1 N+1,\ldots,N\}$ have an equal amount of filled capacity of $C\left(1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^{\hat{\alpha}_1 N}\right)$, based on the analysis in the above paragraph (i.e., Equation (15)).[40] For the subsequent batches of volunteers, i.e., for $j \in \{\hat{\alpha}_1 N+1,\ldots,\hat{\alpha}_2 N\}$, MSVV will maintain an equal fill rate among all compatible opportunities by evenly distributing the $\Delta_j = C$ arriving volunteers in batch j among the $N-j+1$ compatible opportunities. Thus, after the final arrival in batch j (which is the last volunteer compatible with opportunity j), opportunity j will either have reached capacity or will have a filled capacity of

$$C\left(1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^{\hat{\alpha}_1 N}\right) + \sum_{j=\hat{\alpha}_1 N+1}^{i} \frac{C}{N-j+1}.$$

To compute the fraction of total capacity filled under MSVV on $\mathcal{I}_1(\beta)$, we then take an average over the fill rate of all opportunities. To that end, we first compute the fill rate for each opportunity in the limit as the number of opportunities approaches infinity.

For $i \in [\hat{\alpha}_1 N]$,

$$\mathrm{FR}_{i,T} \quad = \quad 1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^{i}$$

Each opportunity $i \in \{\hat{\alpha}_1 N+1,\ldots,\hat{\alpha}_2 N\}$ will not reach capacity, and thus its fill rate approaches:

$$\begin{aligned}
\mathrm{FR}_{i,T} \quad &= \quad 1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^{\hat{\alpha}_1 N} + \sum_{j=\hat{\alpha}_1 N+1}^{i} \frac{1}{N-j+1} \\
&= \quad 1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^{\hat{\alpha}_1 N} + \sum_{k=N-i+1}^{(1-\hat{\alpha}_1)N} \frac{1}{k}
\end{aligned}$$

It is easy to verify algebraically that for $i = \hat{\alpha}_2 N$, the fill rate of opportunity i, $\mathrm{FR}_{i,T}$, asymptotically approaches 1. The remaining opportunities reach capacity.

With this in mind, the fraction of filled capacity under MSVV can be computed as follows:

$$\begin{aligned}
\frac{1}{N}\sum_{i\in[N]} \mathrm{FR}_{i,T} =& \frac{1}{N}\left(\sum_{i=1}^{\hat{\alpha}_1 N}\left(1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^{i}\right)\right. \\
&\left. + \sum_{i=\hat{\alpha}_1 N+1}^{\hat{\alpha}_2 N}\left(1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^{\hat{\alpha}_1 N} + \sum_{k=N-i+1}^{(1-\hat{\alpha}_1)N}\frac{1}{k}\right) + \sum_{i=\hat{\alpha}_2 N+1}^{N} 1\right) \\
\xrightarrow{N\to\infty}& \int_0^{\hat{\alpha}_1} 1-\exp\left(\frac{-x}{1-\hat{\alpha}_1}\right)\,\partial x + \int_{\hat{\alpha}_1}^{\hat{\alpha}_2} 1-\exp\left(\frac{-\hat{\alpha}_1}{1-\hat{\alpha}_1}\right) + \log\left(\frac{1-\hat{\alpha}_1}{1-x}\right)\,\partial x + (1-\hat{\alpha}_2) \qquad (16)
\end{aligned}$$

[40] We allow C to be sufficiently large such that there is vanishing integrality gap.

$$
\begin{aligned}
=&\hat{\alpha}_1-(1-\hat{\alpha}_1)\left(1-\exp\left(\frac{-\hat{\alpha}_1}{1-\hat{\alpha}_1}\right)\right)+(\hat{\alpha}_2-\hat{\alpha}_1)\left[1-\exp\left(\frac{-\hat{\alpha}_1}{1-\hat{\alpha}_1}\right)\right]\\
&\qquad+\int_{\hat{\alpha}_1}^{\hat{\alpha}_2}\log\left(\frac{1-\hat{\alpha}_1}{1-x}\right)\,\partial x+(1-\hat{\alpha}_2)\\
=&(1-\hat{\alpha}_2)\Big(\exp\big(-\hat{\alpha}_1/(1-\hat{\alpha}_1)\big)+\log\big((1-\hat{\alpha}_2)/(1-\hat{\alpha}_1)\big)\Big)+\hat{\alpha}_2\\
=&\hat{\alpha}_2
\end{aligned}
$$

In (16), we again use the fact that $(1-1/n)^{nx}$ approaches e^{-x} as n approaches infinity. Furthermore, we use the fact that $\sum_{k=yn}^{xn} 1/k$ approaches $\log(x/y)$ as n approaches infinity. The last equality comes from applying the definition of $\hat{\alpha}_2$ to see that $\log\big((1-\hat{\alpha}_2)/(1-\hat{\alpha}_1)\big)=-\exp\big(-\hat{\alpha}_1/(1-\hat{\alpha}_1)\big)$. This completes the proof of Claim 1. □

Claim 2 *For any EFET β,* `OPT` *fills all capacity on $\mathcal{I}_1(\beta)$.*

Proof of Claim 2 Consider a solution which matches all external traffic and then matches each of the Δ_i internal traffic volunteers in batch i to opportunity i. To see why such a solution gives a perfect matching, note that each opportunity $i\in\{1,\dots,\hat{\alpha}_1 N\}$ will receive $C\left(1-\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^i\right)$ matches from external traffic and $\Delta_i=C\left(\frac{(1-\hat{\alpha}_1)N}{(1-\hat{\alpha}_1)N+1}\right)^i$ matches from internal traffic, leading to a total of C matches. Each opportunity $i\in\{\hat{\alpha}_1 N+1,\dots,N\}$ will receive $\Delta_i=C$ matches (all from internal traffic). Thus, each opportunity is filled to capacity under this solution, which implies that the optimal solution must also fill all capacity. □

Combining Claims 1 and 2, we see that `MSVV` only fills a fraction $\hat{\alpha}_2$ of the capacity filled by `OPT` on this family of instances, which provides a parameterized upper bound on the competitive ratio of `MSVV` in this setting.[41]

A.3. Proof of Proposition 3 (Section 4.1)

We prove Proposition 3 in two steps. In **Step (a)**, fixing an instance $\mathcal{I}\in\mathcal{I}_\beta$, we show that the expected fraction of capacity filled by external traffic is β under both `AC` and `OPT`. Then, in **Step (b)** we establish a lower bound on the amount of capacity filled by internal traffic under `AC`, which depends on the amount of capacity filled by internal traffic under `OPT`. Together, these steps enable us to place a lower bound on the competitive ratio of `AC`, where the bound is parameterized by the EFET β.

Step (a): Both `OPT` and `AC` always recommend the targeted opportunity i_t^* to external traffic. Since external traffic is assumed to arrive before all internal traffic, this external traffic will fill a fraction of capacity given by

$$
\beta(\mathcal{I})=\frac{\sum_{i\in[n]}\mathbb{E}\left[\min\{c_i,\sum_{t\in\mathcal{V}^{\text{EXT}}}\mathbb{1}[\xi_t(S_t^{i_t^*})=i]\}\right]}{\sum_{i\in[n]}c_i}.
$$

We note that this fraction of capacity is exactly equivalent to the definition of the EFET (see Definition 2), which is equal to β for any instance $\mathcal{I}\in\mathcal{I}_\beta$.

[41] To show that this upper bound holds for any minimum capacity $\underline{c}$, it suffices to add an additional opportunity with capacity $\underline{c}$ for which volunteers have conversion probability of 0. The performance of both `OPT` and `MSVV` are unchanged, and the EFET remains the same in the limit as N approaches infinity.

Step (b): Fixing an instance $\mathcal{I}$, we now turn our attention to lower-bounding the expected amount of capacity filled by internal traffic under the AC algorithm, where the expectation is taken over *sample paths* $\boldsymbol{\omega} = \{\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_T\}$, i.e., realizations of random variables that govern volunteer choices in this instance. Formally, we interpret $\boldsymbol{\omega}_t$ as a vector of length $n+1$, where the i^{th} component of $\boldsymbol{\omega}_t$ (denoted $\omega_{i,t}$) indicates volunteer t's sign-up decision if the platform were to recommend opportunity i. For a fixed instance $\mathcal{I}$ and a fixed sample path $\boldsymbol{\omega}$, we use $\widehat{\texttt{AC}}$ to denote the amount of capacity filled by internal traffic under the AC algorithm.[42]

Our lower bound on $\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{AC}}\right]$ will depend on the expected amount of capacity filled by internal traffic under OPT, which we likewise denote with $\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right]$. Note that in this step of the proof, we are concerned only with the remaining capacities for each opportunity i after the arrival of external traffic (denoted $\hat{c}_i$), which depends on the realizations of sign-ups made by external traffic. As such, $\hat{c}_i$ depends not only on the instance $\mathcal{I}$, but also on the sample path $\boldsymbol{\omega}$.

To provide such a lower bound, we leverage the LP-free approach developed in Goyal and Udwani (2019) and Goyal et al. (2020), which involves the creation of path-based pseudo-rewards. (For a more complete discussion of the intuition behind this approach, we kindly refer to the proof sketch of Theorem 2 in Section 4.4.) For a fixed instance $\mathcal{I}$ and a fixed sample path $\boldsymbol{\omega}$, we define the pseudo-rewards $\hat{L}_t$ for all $t \in \mathcal{V}^{\text{INT}}$ and $\hat{K}_i$ for all $i \in [n]$ according to the following:

$$\hat{L}_t = \sum_{i \in [n]} \psi(\text{FR}_{i,t-1}) \mathbb{1}[\xi_t(S_t^{\texttt{OPT}}) = i] \tag{17}$$

$$\hat{K}_i = \sum_{t \in \mathcal{V}^{\text{INT}}} (1 - \psi(\text{FR}_{i,t-1}))\, \mathbb{1}[\xi_t(S_t^{\texttt{AC}}) = i], \tag{18}$$

where we remind that under the AC algorithm, $\text{FR}_{i,t-1} = \texttt{AC}_{i,t}^{\text{INT}} / (c_i - \texttt{AC}_{i,t}^{\text{EXT}})$. This is equivalent to $\texttt{AC}_{i,t}^{\text{INT}} / \hat{c}_i$ in our warm-up setting where external traffic arrives first and the remaining capacity for opportunity i is given by $\hat{c}_i$. We now prove that the expected sum of these pseudo-rewards serves as a lower bound on the expected value of $\widehat{\texttt{AC}}$.

Lemma 7 *For any instance $\mathcal{I}$,*

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{AC}}\right] \quad \geq \quad \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t \in \mathcal{V}^{\text{INT}}} \hat{L}_t + \sum_{i \in [n]} \hat{K}_i\right], \tag{19}$$

where $\hat{L}_t$ and $\hat{K}_i$ are defined in (17) and (18), respectively.

Proof of Lemma 7: The proof follows from the definition of $\hat{L}_t$ and $\hat{K}_i$ as well as the design of the AC algorithm:

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{AC}}\right] = \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t \in \mathcal{V}^{\text{INT}}} \sum_{i \in [n]} \mathbb{1}[\xi_t(S_t^{\texttt{AC}}) = i]\right] \tag{20}$$

[42] Even though $\widehat{\texttt{AC}}$ depends on the instance and the sample path, we hereafter suppress this dependence to ease exposition (for $\widehat{\texttt{AC}}$ as well as for all other quantities that depend on the instance and the sample path).

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{INT}}}\sum_{i\in[n]}\psi(\text{FR}_{i,t-1})\mathbb{1}[\xi_t(S_t^{\texttt{AC}})=i] + \sum_{t\in\mathcal{V}^{\text{INT}}}\sum_{i\in[n]}\Big(1-\psi(\text{FR}_{i,t-1})\Big)\mathbb{1}[\xi_t(S_t^{\texttt{AC}})=i]\right] \tag{21}$$

$$\geq \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{INT}}}\sum_{i\in[n]}\psi(\text{FR}_{i,t-1})\mathbb{1}[\xi_t(S_t^{\texttt{OPT}})=i] + \sum_{i\in[n]}\hat{K}_i\right] \tag{22}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{INT}}}\hat{L}_t + \sum_{i\in[n]}\hat{K}_i\right] \tag{23}$$

Equality (20) holds because the `AC` algorithm will never recommend an opportunity that has already reached capacity to internal traffic.[43] Consequently, the amount of capacity filled by internal traffic under the `AC` algorithm is exactly equal to the numbers of sign-ups from internal traffic.

Inequality (22) follows from the `AC` algorithm's optimality condition (see Algorithm 2), which ensures that it recommends the opportunity that maximizes the weighted probability of generating a sign-up (where the weight for opportunity i at time t is given by $\psi(\text{FR}_{i,t-1})$). Since the recommendation provided by `OPT` to any volunteer must be independent of their sign-up realization, the inequality holds. Applying the definition of the pseudo-rewards $\hat{L}_t$ for $t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0$ completes the proof of Lemma 7. □

Next, we place a lower bound on the expected sum of the pseudo-rewards, which depends on the amount of capacity of each opportunity i filled by internal traffic under `OPT` along a fixed sample path, which we denote by $\widehat{\texttt{OPT}}_i$.

Lemma 8 *For any instance $\mathcal{I}$,*

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{INT}}}\hat{L}_t + \sum_{i\in[n]}\hat{K}_i\right] \quad\geq\quad (1-1/e)\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right] - \sum_{i\in[n]}\mathbb{E}_{\boldsymbol{\omega}}\left[\mathbb{1}\left[\widehat{\texttt{OPT}}_i = \hat{c}_i\right]\right] \tag{24}$$

Proof of Lemma 8: We will prove this claim along each sample path $\boldsymbol{\omega}$ by separately placing lower bounds on the $\hat{L}_t$ pseudo-rewards and the $\hat{K}_i$ pseudo-rewards. For the former,

$$\sum_{t\in\mathcal{V}^{\text{INT}}}\hat{L}_t \quad=\quad \sum_{t\in\mathcal{V}^{\text{INT}}}\sum_{i\in[n]}\psi(\text{FR}_{i,t-1})\mathbb{1}[\xi_t(S_t^{\texttt{OPT}})=i] \tag{25}$$

$$\geq\quad \sum_{t\in\mathcal{V}^{\text{INT}}}\sum_{i\in[n]}\psi(\text{FR}_{i,T})\mathbb{1}[\xi_t(S_t^{\texttt{OPT}})=i] \tag{26}$$

$$=\quad \sum_{i\in[n]}\widehat{\texttt{OPT}}_i\psi\left(\text{FR}_{i,T}\right) \tag{27}$$

Equality in (25) follows from the definition of $\hat{L}_t$. Inequality in (26) holds because ψ is a decreasing function in its argument, and $\text{FR}_{i,T}\geq\text{FR}_{i,t-1}$ for all $t\in[T]$. All other steps are algebraic.

We now turn our attention to the $\hat{K}_i$ pseudo-rewards:

$$\hat{K}_i \quad=\quad \sum_{t\in\mathcal{V}^{\text{INT}}}(1-\psi(\text{FR}_{i,t-1}))\,\mathbb{1}[\xi_t(S_t^{\texttt{AC}})=i] \tag{28}$$

[43] To see this, note that if opportunity i has reached capacity before time t, then $\mu_{i,t}\cdot\psi(\text{FR}_{i,t-1})=0$. Based on its convention for breaking ties in favor of the opportunity with the lowest index, the `AC` algorithm would always recommend opportunity 0 instead of an at-capacity opportunity i.

$$= \sum_{t\in\mathcal{V}^{\text{INT}}} \left(1-\psi\left(\frac{\texttt{AC}^{\text{INT}}_{i,t-1}}{\hat{c}_i}\right)\right)\mathbb{1}[\xi_t(S^{\texttt{AC}}_t)=i] \tag{29}$$

$$= \sum_{k\in[\texttt{AC}^{\text{INT}}_{i,T}]} \left(1-\psi\left(\frac{k-1}{\hat{c}_i}\right)\right) \tag{30}$$

$$= e^{\frac{-1}{\hat{c}_i}} \sum_{k\in[\texttt{AC}^{\text{INT}}_{i,T}]} \left(1-\psi\left(\frac{k}{\hat{c}_i}\right)\right) \tag{31}$$

$$\geq e^{\frac{-1}{\hat{c}_i}} \int_0^{\texttt{AC}^{\text{INT}}_{i,T}} 1-\psi(x/\hat{c}_i)\ \partial x \tag{32}$$

$$= e^{\frac{-1}{\hat{c}_i}}\hat{c}_i\left(1-\psi\left(\text{FR}_{i,T}\right)-1/e\right) \tag{33}$$

$$\geq (\hat{c}_i-1)\left(1-\psi\left(\text{FR}_{i,T}\right)-1/e\right) \tag{34}$$

Equality in (30) holds because the counter $\texttt{AC}^{\text{INT}}_{i,t}$ will increase by 1 for any $t\in\mathcal{V}^{\text{INT}}$ where $\xi_t(S^{\texttt{AC}}_t)=i$. The summation in (30) represents a left Reimann sum of an increasing function. In (31), we utilize the fact that for any k, $1-\psi((k-1)/\hat{c}_i)=e^{-1/\hat{c}_i}(1-\psi(k/\hat{c}_i))$. As the summation in (31) is now a right Reimann sum of an increasing function, we bound the sum with an appropriate integral in (32). Finally, (34) holds because $e^{-x}\geq 1-x$ for any x.

Combining (27) and (34), we see that for each sample path $\boldsymbol{\omega}$,

$$\begin{aligned}\sum_{t\in\mathcal{V}^{\text{INT}}}\hat{L}_t+\sum_{i\in[n]}\hat{K}_i &\geq \sum_{i\in[n]}\widehat{\texttt{OPT}}_i\psi\left(\text{FR}_{i,T}\right)+(\hat{c}_i-1)\left(1-\psi\left(\text{FR}_{i,T}\right)-1/e\right) && (35)\\ &\geq \sum_{i\in[n]}\left(\widehat{\texttt{OPT}}_i-\mathbb{1}[\widehat{\texttt{OPT}}_i=\hat{c}_i]\right)\psi\left(\text{FR}_{i,T}\right)+\left(\widehat{\texttt{OPT}}_i-\mathbb{1}[\widehat{\texttt{OPT}}_i=\hat{c}_i]\right)\left(1-\psi\left(\text{FR}_{i,T}\right)-1/e\right) && (36)\\ &=(1-1/e)\sum_{i\in[n]}\left(\widehat{\texttt{OPT}}_i-\mathbb{1}\left[\widehat{\texttt{OPT}}_i=\hat{c}_i\right]\right) && (37)\\ &\geq(1-1/e)\cdot\widehat{\texttt{OPT}}-\sum_{i\in[n]}\mathbb{1}\left[\widehat{\texttt{OPT}}_i=\hat{c}_i\right] && (38)\end{aligned}$$

Inequality in (36) comes from noting that $\left(\widehat{\texttt{OPT}}_i-\mathbb{1}\left[\widehat{\texttt{OPT}}_i=\hat{c}_i\right]\right)$ cannot exceed either $\widehat{\texttt{OPT}}_i$ or $\hat{c}_i-1$. (We note that the binary indicator $\mathbb{1}\left[\widehat{\texttt{OPT}}_i=\hat{c}_i\right]$ is equal to 1 if and only if opportunity i reaches capacity under $\texttt{OPT}$ along the fixed sample path $\boldsymbol{\omega}$.) Taking expectation across all sample paths completes the proof of Lemma 8. □

Combining Lemmas 7 and 8, we see that we can bound the expected amount of capacity filled by internal traffic under $\texttt{AC}$ via the following inequality:

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{AC}}\right] \quad\geq\quad (1-1/e)\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right]-\sum_{i\in[n]}\mathbb{E}_{\boldsymbol{\omega}}\left[\mathbb{1}\left[\widehat{\texttt{OPT}}_i=\hat{c}_i\right]\right] \tag{39}$$

Together with Step (a), we have shown that for any instance $\mathcal{I}\in\mathcal{I}_\beta$,

$$\begin{aligned}\frac{\mathbb{E}_{\boldsymbol{\omega}}\left[\texttt{AC}\right]}{\mathbb{E}_{\boldsymbol{\omega}}\left[\texttt{OPT}\right]} &= \frac{\beta\cdot\sum_{i\in[n]}c_i+\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{AC}}\right]}{\beta\cdot\sum_{i\in[n]}c_i+\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right]} && (40)\\ &\geq \frac{\beta\cdot\sum_{i\in[n]}c_i+(1-1/e)\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right]-\sum_{i\in[n]}\mathbb{E}_{\boldsymbol{\omega}}\left[\mathbb{1}\left[\widehat{\texttt{OPT}}_i=\hat{c}_i\right]\right]}{\beta\cdot\sum_{i\in[n]}c_i+\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right]} && (41)\end{aligned}$$

$$= \frac{\beta \cdot \sum_{i\in[n]} c_i + (1-1/e)\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right]}{\beta \cdot \sum_{i\in[n]} c_i + \mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right]} - \frac{\sum_{i\in[n]} \mathbb{E}_{\boldsymbol{\omega}}\left[\mathbb{1}\left[\widehat{\texttt{OPT}}_i = \hat{c}_i\right]\right]}{\mathbb{E}_{\boldsymbol{\omega}}\left[\texttt{OPT}\right]} \tag{42}$$

$$\geq \frac{\beta \cdot \sum_{i\in[n]} c_i + (1-1/e)\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right]}{\beta \cdot \sum_{i\in[n]} c_i + \mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right]} - \underline{c}^{-1} \tag{43}$$

$$\geq \beta + (1-\beta)(1-1/e) - \underline{c}^{-1} \tag{44}$$

Equality in (40) comes from applying the result of Step (a), while inequality in (41) comes from applying the result of Step (b). Equality in (41) follows from the definition of $\texttt{OPT}$. To see that (43) holds, we first fix a sample path. Along that sample path, if $\widehat{\texttt{OPT}}_i = \hat{c}_i$, then opportunity i must have reached capacity under $\texttt{OPT}$. The capacity of opportunity i is at least $\underline{c}$. Thus, along every sample path, $\texttt{OPT} \geq \underline{c} \sum_{i\in[n]} \mathbb{E}_{\boldsymbol{\omega}}\left[\mathbb{1}\left[\widehat{\texttt{OPT}}_i = \hat{c}_i\right]\right]$. This is a sufficient condition to establish (43).

Finally, (44) comes from noting that the expression in (43) is decreasing in $\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right]$, which can be at most $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} \hat{c}_i\right]$. Furthermore, $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} \hat{c}_i\right] = (1-\beta)\sum_{i\in[n]} c_i$. We then plug in this upper bound for $\mathbb{E}_{\boldsymbol{\omega}}\left[\widehat{\texttt{OPT}}\right]$. This final inequality establishes a lower bound for any instance $\mathcal{I} \in \mathcal{I}_\beta$. Thus, it represents a lower bound on the competitive ratio parameterized by the EFET β, as desired. This completes the proof of Proposition 3. □

A.4. Proof of Theorem 1 (Section 4.2)

This proof is an adaptation of the proof of Theorem 7.1 in Mehta et al. (2007), which we have generalized to apply in our setting. We aim to prove that no online algorithm (deterministic or randomized) can provide a competitive ratio greater than $\max\{1-1/e, 1+\beta\log(\beta)\}$. By Yao's lemma (Yao 1977), it is sufficient to show that there exists a distribution over a set of instances for which no deterministic algorithm can provide an expected value greater than $\max\{1-1/e, 1+\beta\log(\beta)\}\texttt{OPT}$.

We begin by fixing an EFET β and describing an instance $\mathcal{I}_2(\beta)$. In this instance, the set of opportunities is of size N, each with identical large capacity C. The arrival sequence consists of NC volunteers. The first $(1-\beta)NC$ of these volunteers are internal traffic, and the remaining βNC are external traffic.[44] All volunteers have conversion probabilities of 1 or 0, and if $\mu_{i,t} = 1$ (resp. 0), we will refer to opportunity i and volunteer t as *compatible* (resp. incompatible).

The arrival sequence of $\mathcal{I}_2(\beta)$ can be broken down into N batches of C sequentially-arriving identical volunteers. For each $j \in \{1, \dots, (1-\beta)N\}$, the j^{th} batch of volunteers consists of internal traffic that is compatible with all opportunities $i \geq j$. For each $j \in \{(1-\beta)N + 1, \dots, N\}$, the j^{th} batch of volunteers consists of external traffic which views (and is compatible with) opportunity $i^*_j = j$. This external traffic can fill the entire capacity of each of these βN opportunities, which implies that the EFET is equal to β in such an instance.

We first establish the value of $\texttt{OPT}$ on instance $\mathcal{I}_2(\beta)$.

Claim 3 *For any EFET β, $\texttt{OPT}$ achieves a value of NC on $\mathcal{I}_2(\beta)$.*

[44] We assume that βN is an integer. This assumption does not impact the upper bound in the statement of Theorem 1, as the expression comes from taking the limit as N and C approach ∞.

Proof of Claim 3: Consider a solution which matches each of the C volunteers in the j^{th} batch to opportunity j, for all $j \in [N]$. These volunteer-opportunity pairs are all compatible based on the compatibility structure previously described, and (since the conversion probabilities are exactly equal to 1) each opportunity will exactly reach its capacity of C. As this solution fills all capacity, `OPT` must also fill all capacity, thereby achieving a total value of NC, regardless of the EFET β. $\square$

We now consider the set of instances which can be obtained from $\mathcal{I}_2(\beta)$ by permuting the indices of the opportunities. Specifically, we apply a permutation $\mathcal{P}$ to the set of opportunities such that any algorithm sees opportunities with indices $\{\mathcal{P}(1),\dots,\mathcal{P}(N)\}$. We highlight that *a priori* the opportunities appear identical to an online algorithm, aside from their indices. We augment our previous notation and describe such an instance as $\mathcal{I}_2(\beta,\mathcal{P})$.

Suppose that the permutation $\mathcal{P}$ is drawn uniformly at random from the set of all permutations of N indices. For the set of instances generated by this distribution over permutations, in the following lemma, we place an upper-bound on the expected value of any deterministic online algorithm.

Claim 4 *Consider any deterministic online algorithm π. For any EFET β,*

$$\mathbb{E}_{\mathcal{P}}[\pi(\mathcal{I}_2(\beta,\mathcal{P}))] \le \sum_{i\in[(1-\beta)N]} \min\left\{C, \sum_{j=1}^{i} \frac{C}{N-j+1}\right\} + \sum_{i\in[N]\setminus[(1-\beta)N]} C, \tag{45}$$

where the expectation is taken with respect to the uniform distribution over permutations $\mathcal{P}$.

Proof of Claim 4: To aid in this proof, let us define $d_{i,j}$ as the amount of volunteers allocated to opportunity i from the j^{th} batch of arriving volunteers. Recall that for $j \in \{1,\dots,(1-\beta)N\}$, volunteers in the j^{th} batch of arrivals are compatible with all opportunities $i \ge j$. Thus, we have:

$$E_{\mathcal{P}}[d_{i,j}] \le \begin{cases} \dfrac{C}{N-j+1}, & \text{if } i \ge j \\ 0 & \text{if } i < j \end{cases}$$

To see why this must be the case, note that for each volunteer in the j^{th} batch of volunteers, there are a total of $N-j+1$ compatible opportunities. The online algorithm cannot distinguish between these compatible opportunities, as it only observes the indices $\{\mathcal{P}(j),\dots,\mathcal{P}(N)\}$. Specifically, if i is one such compatible opportunity, the online algorithm does not know which index in the set $\{\mathcal{P}(j),\dots,\mathcal{P}(N)\}$ is equal to $\mathcal{P}(i)$. Hence, the expected amount of volunteers allocated to opportunity i cannot exceed $\frac{C}{N-j+1}$, when taking expectation with respect to the uniform distribution over permutations $\mathcal{P}$.

More simply, for batches $j \in \{(1-\beta)N+1, N\}$, the volunteers are only compatible with opportunity i if $i=j$. Hence,

$$E_{\mathcal{P}}[d_{i,j}] \le \begin{cases} C & \text{if } i = j \\ 0 & \text{if } i \ne j \end{cases}$$

After the arrival of all volunteers, the expected fill of opportunity i is upper-bounded by $\sum_{j\in[N]} E_{\mathcal{P}}[d_{i,j}]$. This quantity is either C (if $i > (1-\beta)N$) or $\min\{C, \sum_{j=1}^{i} \frac{C}{N-j+1}\}$ (if $i \le (1-\beta)N$). Summing over all opportunities, we have the following upper bound on the value of any online algorithm:

$$\sum_{i\in[(1-\beta)N]} \min\left\{C, \sum_{j=1}^{i} \frac{C}{N-j+1}\right\} + \sum_{i\in[N]\setminus[(1-\beta)N]} C \tag{46}$$

This completes the proof of Claim 4. $\square$

Together, and in combination with Yao's lemma, these claims establish an upper-bound on the achievable competitive ratio of any online algorithm of

$$\frac{1}{NC}\left(\sum_{i\in[(1-\beta)N]}\min\left\{C,\sum_{j=1}^{i}\frac{C}{N-j+1}\right\}+\sum_{i\in[N]\setminus[(1-\beta)N]}C\right)\quad\rightarrow\quad\max\{1-1/e,1+\beta\log(\beta)\},\tag{47}$$

where the limit holds as C and N approach infinity.[45] □

A.5. Upper Bound on MSVV in General Settings (Section 4.3)

In the following proposition, we provide an upper bound on the competitive ratio of MSVV as a function of the EFET β.

Proposition 6 (Upper Bound on MSVV) *For any effective fraction of external traffic β and any minimum capacity,* MSVV *cannot achieve a competitive ratio better than*

$$\begin{cases}1-1/e, & \beta\le 1/e\\ \min\{\alpha_2,\alpha_3\}, & \beta>1/e\end{cases}$$

where, for $\beta>1/e$, α_2 is given by

$$\alpha_2=1-\frac{1-\alpha_1}{\exp\left(\exp(-\alpha_1/(1-\alpha_1))\right)}$$

and α_1 and is the unique solution in $[0,1]$ to $\beta=\alpha_1+(1-\alpha_1)\Big(\exp\big(-\alpha_1/(1-\alpha_1)\big)-1\Big)+\frac{1-\alpha_1}{\exp(\exp(-\alpha_1/(1-\alpha_1)))}$. In addition,

$$\alpha_3=\min_{\alpha_4\in[0,\beta]}1-\frac{1-\beta}{1-\alpha_4}\left(\alpha_5+(1-\alpha_6)\log\left(\frac{1-\alpha_5}{1-\alpha_6}\right)\right),$$

where $\alpha_5=\min\{1-\alpha_4,\frac{\alpha_4(\beta-\alpha_4)}{1-\beta}\}$ and $\alpha_6=\min\{1-\alpha_4,1-(1-\alpha_5)/e\}$.

Proof of Proposition 6 The first part of Proposition 6 – which establishes an upper bound of $1-1/e$ when $\beta\le 1/e$ – follows immediately from Theorem 1, in which we prove such an upper bound on the competitive ratio of *any* online algorithm.

We prove the remainder of this proposition in two claims by showing two different upper bounds (α_2 and α_3) on the competitive ratio of MSVV parameterized by the effective fraction of external traffic β. Proposition 6 follows by taking the minimum of the two upper bounds for a given β.

To prove each claim, we construct a family of instances parameterized by β. We then evaluate the value of MSVV on that family of instances relative to the value of OPT. Both of the instances that we design leverage the fact that the notion of a fill rate under MSVV does not distinguish between internal and external traffic. As a result, MSVV may mistakenly withhold internal traffic from opportunities that have previously received external traffic. Furthermore, all instances leverage the triangular structure of our general hardness result (see Appendix A.4).

---

[45] To show that this upper bound holds for any minimum capacity $\underline{c}$, it suffices to add an additional opportunity with capacity $\underline{c}$ for which volunteers have conversion probability of 0. The value of OPT and the upper bound on the performance of any algorithm do not change, and the EFET also remains the same in the limit as N approaches infinity.

Claim 5 *For any effective fraction of external traffic $\beta \in (1/e, 1]$, the competitive ratio of* MSVV *is at most*

$$1 - \frac{1-\alpha_1}{\exp\left(\exp(-\alpha_1/(1-\alpha_1))\right)} \tag{48}$$

where α_1 is the unique solution in $[0,1]$ to $\beta = \alpha_1 + (1-\alpha_1)\Big(\exp\big(-\alpha_1/(1-\alpha_1)\big) - 1\Big) + \frac{1-\alpha_1}{\exp(\exp(-\alpha_1/(1-\alpha_1)))}$.

Proof of Claim 5 To prove this claim, we construct a family of instances $\mathcal{I}_3(\beta)$ parameterized by the EFET β. (As we will highlight below, this family of instances will have a close relationship to the family of instances $\mathcal{I}_1(\beta)$, introduced in the proof of Proposition 2.) In each instance, there are a large number of opportunities N, each with identical large capacity C. The arrival sequence consists of NC volunteers, and for a given effective fraction of external traffic β, the first βNC of these volunteers are external traffic.[46] All volunteers have conversion probabilities of 1 or 0, and if $\mu_{i,t} = 1$ (resp. 0), we will refer to opportunity i and volunteer t as *compatible* (resp. incompatible).

To help describe the compatibility structure of the arriving volunteers, we first define constants α_1 and α_2. For $\beta \leq 1/e$, we define $\alpha_1 = 0$, while for $\beta > 1/e$, we define α_1 as the unique solution in $[0,1]$[47] to

$$\beta = \alpha_1 + (1-\alpha_1)\Big(\exp\big(-\alpha_1/(1-\alpha_1)\big) - 1\Big) + \frac{1-\alpha_1}{\exp\left(\exp(-\alpha_1/(1-\alpha_1))\right)},$$

and α_2 is defined as

$$\alpha_2 = 1 - \frac{1-\alpha_1}{\exp\left(\exp(-\alpha_1/(1-\alpha_1))\right)}.$$

The arrival sequence begins with external traffic volunteers for the first $\alpha_1 N$ opportunities. Specifically, for each opportunity $i \in \{1, \ldots, \alpha_1 N\}$, there are $C\left(1 - \left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^i\right)$ compatible external traffic arrivals for that opportunity. After the arrival of these volunteers, the internal traffic arrives, according to the following compatibility structure: for each opportunity $i \in \{1, \ldots, \alpha_2 N\}$, there is a batch of Δ_i sequentially-arriving homogeneous volunteers. The batches consist of $\Delta_i = C\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^i$ volunteers for each $i \in \{1, \ldots, \alpha_1 N\}$, and they consist of $\Delta_i = C$ volunteers for each $i \in \{\alpha_1 N + 1, \ldots, \alpha_2 N\}$. Opportunities in batch i are compatible with all opportunities $j \geq i$. Finally, the arrival sequence concludes with $(1-\alpha_2)N$ batches of C external traffic volunteers, where each batch views (and is compatible with) one opportunity $i \in \{\alpha_2 N + 1, \ldots, N\}$.

Before analyzing this family of instances, we make two observations. First, this arrival sequence is quite similar to the arrival sequence in the family of instances $\mathcal{I}_1(\beta)$, which are visualized in Figure 6 and which provide our upper bound on MSVV in the setting where all external traffic arrives first (see Proposition 2). The only difference comes from the last batches of arrivals, which are external traffic in this family of instances (as opposed to internal traffic with broader compatibility, as in $\mathcal{I}_1(\beta)$). In both cases, these volunteers are unable to be allocated under MSVV as their compatible opportunities have already reached capacity, whereas these volunteers are allocated under OPT. Hence, the value of MSVV and the value of OPT are both unchanged. Crucially, though, the EFET is different in these two instances, due to the change in source of the last-arriving

[46] We assume that $(1-\beta)NC$ is an integer. This assumption does not impact the upper bound in the statement of Claim 5, as the expression comes from taking the limit as N approaches ∞.

[47] We note that for any $\beta \in (1/e, 1]$, it is easy to verify numerically that there is a unique solution in the interval $[0,1]$ for α_1.

volunteers. As a result, for a fixed β, the instance $\mathcal{I}_1(\beta)$ and $\mathcal{I}_3(\beta)$ differ significantly. Instead, $\mathcal{I}_1(\beta)$ and $\mathcal{I}_3(\beta+\hat{\alpha}_2)$ are nearly identical (where $\hat{\alpha}_2$ is a function of β, as defined in the proof of Proposition 2). This relationship means the upper bound provided by the family of instances $\mathcal{I}_3(\beta)$ is a non-linear transformation of the upper bound provided by the family of instances $\mathcal{I}_1(\beta)$. Furthermore, we remark that in the limit as β approaches $1/e$, $\mathcal{I}_3(\beta)$ approaches the instance $\mathcal{I}_2(1/e)$, which provides our general hardness result presented in Theorem 1.

We now verify that the EFET is equal to β in the limit as N gets large.

$$
\begin{aligned}
\frac{1}{NC}\left(\sum_{i=1}^{\alpha_1 N} C\left(1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^i\right)+(1-\alpha_2)NC\right) &= \frac{1}{N}\left[\sum_{i=1}^{\alpha_1 N}\left(1-\left(1-\frac{1}{(1-\alpha_1)N+1}\right)^i\right)+(1-\alpha_2)N\right] \\
&\xrightarrow{N\to\infty} \int_0^{\alpha_1}\left[1-\exp\left(\frac{-x}{1-\alpha_1}\right)\,\partial x\right]+(1-\alpha_2) \qquad (49) \\
&= \left(\alpha_1+(1-\alpha_1)\left(\exp\left(\frac{-\alpha_1}{1-\alpha_1}\right)-1\right)\right)+(1-\alpha_2) \\
&= \beta \qquad (50)
\end{aligned}
$$

In (49), we use the fact that $(1-1/n)^{nx}$ approaches e^{-x} as n approaches infinity. Furthermore, (50) follows by applying the definitions of α_2 and α_1. Next, we analyze the value of `MSVV` and `OPT` on the above family of instances.

Value of `MSVV` on Instance $\mathcal{I}_3(\beta)$: We will show that for any $EFET$ β, the fraction of total capacity filled under `MSVV` on $\mathcal{I}_3(\beta)$ is at most α_2. To that end, we will first bound the amount of filled capacity for each opportunity under `MSVV`. First, we will show that the $\alpha_1 N$ opportunities that initially receive external traffic do not receive any matches from internal traffic; i.e., for each $i\in[\alpha_1 N]$, we will show that $\texttt{MSVV}_{i,T} = C\left(1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^i\right)$. Suppose towards a contradiction that there exists some opportunity $j\in[\alpha_1 N]$ which receives a match from internal traffic under `MSVV`. Due to restrictions on compatibility, this match must have come from one of the first j batches of internal traffic, which in total represents

$$
\sum_{i=1}^{j}\Delta_i = \sum_{i=1}^{j} C\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^i = C\left((1-\alpha_1)N\right)\left(1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^j\right)
$$

internal traffic volunteers. We are supposing that one of these volunteers was allocated to opportunity j. In that case, due to the pigeonhole principle, there must be at least one opportunity j' – from among the $(1-\alpha_1)N$ opportunities that *did not* initially receive external traffic – with a filled capacity strictly less than $C\left(1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^j\right)$ upon the arrival of the last volunteer in batch j. By definition, `MSVV` should never have recommended j ahead of j', giving us a contradiction.

Next, we show that each opportunity $i\in\{\alpha_1 N+1,\dots,\alpha_2 N\}$ has a filled capacity of

$$
\texttt{MSVV}_{i,T} \le C\left(1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^{\alpha_1 N}+\sum_{j=\alpha_1 N+1}^{i}\frac{1}{N-j+1}\right). \qquad (51)
$$

Note that in this matching setting, `MSVV` recommends opportunities to equalize their fill rate. Thus, after the arrival of the $\alpha_1 N^{\text{th}}$ batch of volunteers, all opportunities $j\in\{\alpha_1 N+1,\dots,\alpha_2 N\}$ have an equal amount of filled capacity of $C\left(1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^{\alpha_1 N}\right)$, based on the analysis in the above paragraph (i.e., (51)).[48] For the subsequent batches of internal traffic volunteers, i.e., for $j\in\{\alpha_1 N+1,\ldots,\alpha_2 N\}$, MSVV will maintain an equal fill rate among all compatible opportunities by evenly distributing the $\Delta_j = C$ arriving volunteers in batch j among the $N-j+1$ compatible opportunities. Thus, after the final arrival in batch j (which is the last volunteer compatible with opportunity j), opportunity j will have a filled capacity of at most

$$C\left(1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^{\alpha_1 N}\right)+\sum_{j=\alpha_1 N+1}^{i}\frac{C}{N-j+1}.$$

The remaining opportunities (i.e., opportunities i for $i>\alpha_2 N$) will, at most, reach capacity.

To compute the fraction of total capacity filled under MSVV on $\mathcal{I}_3(\beta)$, we then take an average over the fill rate of all opportunities. To that end, we first compute the fill rate for each opportunity in the limit as the number of opportunities approaches infinity.

For $i\in[\alpha_1 N]$,

$$\mathrm{FR}_{i,T} \quad = \quad 1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^{i}$$

Each opportunity $i\in\{\alpha_1 N+1,\ldots,\alpha_2 N\}$ has a fill rate which is bounded by:

$$\begin{aligned}\mathrm{FR}_{i,T} \quad &= \quad 1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^{\alpha_1 N}+\sum_{j=\alpha_1 N+1}^{i}\frac{1}{N-j+1}\\ &= \quad 1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^{\alpha_1 N}+\sum_{k=N-i+1}^{(1-\alpha_1)N}\frac{1}{k}\end{aligned}$$

It is easy to verify algebraically that for $i=\alpha_2 N$, the fill rate of opportunity i, $\mathrm{FR}_{i,T}$, asymptotically approaches 1. The remaining opportunities reach capacity.

With this in mind, the fraction of filled capacity under MSVV can be computed as follows:

$$\begin{aligned}\frac{1}{N}\sum_{i\in[N]}\mathrm{FR}_{i,T} &= \frac{1}{N}\left(\sum_{i=1}^{\alpha_1 N}\left(1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^{i}\right)\right.\\ &\qquad\left.+\sum_{i=\alpha_1 N+1}^{\alpha_2 N}\left(1-\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^{\alpha_1 N}+\sum_{k=N-i+1}^{(1-\alpha_1)N}\frac{1}{k}\right)+\sum_{i=\alpha_2 N+1}^{N}1\right)\\ &\xrightarrow{N\to\infty}\int_0^{\alpha_1}1-\exp\left(\frac{-x}{1-\alpha_1}\right)\ \partial x+\int_{\alpha_1}^{\alpha_2}1-\exp\left(\frac{-\alpha_1}{1-\alpha_1}\right)+\log\left(\frac{1-\alpha_1}{1-x}\right)\ \partial x+(1-\alpha_2) \qquad (52)\\ &=\alpha_1-(1-\alpha_1)\left(1-\exp\left(\frac{-\alpha_1}{1-\alpha_1}\right)\right)+(\alpha_2-\alpha_1)\left[1-\exp\left(\frac{-\alpha_1}{1-\alpha_1}\right)\right]\\ &\qquad+\int_{\alpha_1}^{\alpha_2}\log\left(\frac{1-\alpha_1}{1-x}\right)\ \partial x+(1-\alpha_2)\\ &=(1-\alpha_2)\Big(\exp\big(-\alpha_1/(1-\alpha_1)\big)+\log\big((1-\alpha_2)/(1-\alpha_1)\big)\Big)+\alpha_2\\ &=\alpha_2\end{aligned}$$

In (52), we again use the fact that $(1-1/n)^{nx}$ approaches e^{-x} as n approaches infinity. Furthermore, we use the fact that $\sum_{k=yn}^{xn}1/k$ approaches $\log(x/y)$ as n approaches infinity. The last equality comes from applying the definition of α_2 to see that $\log\big((1-\alpha_2)/(1-\alpha_1)\big)=-\exp\big(-\alpha_1/(1-\alpha_1)\big)$.

[48] We allow C to be sufficiently large such that there is vanishing integrality gap.

Value of `OPT` on Instance $\mathcal{I}_3(\beta)$: We next show that for any EFET $\beta \in [1/e, 1]$, `OPT` fills all capacity on $\mathcal{I}_3(\beta)$. To see this, consider a solution which matches all external traffic and matches each of the Δ_i internal traffic volunteers in batch i to opportunity i. To see why such a solution gives a perfect matching, note that each opportunity $i \in \{1, \ldots, \alpha_1 N\}$ will receive $C\left(1 - \left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^i\right)$ matches from external traffic and $\Delta_i = C\left(\frac{(1-\alpha_1)N}{(1-\alpha_1)N+1}\right)^i$ matches from internal traffic, leading to a total of C matches. Each opportunity $i \in \{\alpha_1 N + 1, \ldots, N\}$ will receive $\Delta_i = C$ matches (either all from internal traffic or all from external traffic). Thus, each opportunity is filled to capacity under this solution, which implies that the optimal algorithm must also fill all capacity.

Combining the upper bound on the fraction of capacity filled by `MSVV` with the fact that `OPT` fills all capacity, we see that `MSVV` only fills a fraction α_2 of the capacity filled by `OPT` on this family of instances. This provides a parameterized upper bound on the competitive ratio of `MSVV`, and thereby completes the proof of Claim 5. $\square$

Claim 6 *For any effective fraction of external traffic $\beta \in [0,1]$, the competitive ratio of* `MSVV` *is at most*

$$\min_{\alpha_4 \in [0,\beta]} 1 - \frac{1-\beta}{1-\alpha_4}\left(\alpha_5 + (1-\alpha_6)\log\left(\frac{1-\alpha_5}{1-\alpha_6}\right)\right) \tag{53}$$

where $\alpha_5 = \min\{1-\alpha_4, \frac{\alpha_4(\beta-\alpha_4)}{1-\beta}\}$ *and* $\alpha_6 = \min\{1-\alpha_4, 1-(1-\alpha_5)/e\}$.

Consider a family of instances $\mathcal{I}_4(\beta)$, parameterized by the EFET β. Each instance has N opportunities, each with identical large capacity C. Each instance in this family is also parameterized by $\alpha_4 \in [0, \beta]$, which separates the N into subsets of size $(1-\alpha_4)N$ and $\alpha_4 N$.[49] These two subsets will receive external traffic at different times: the former subset will receive external traffic at the beginning of the arrival sequence, while the latter will receive external traffic at the end of the arrival sequence. The full arrival sequence consists of NC volunteers, all of whom have conversion probabilities of 1 or 0. If $\mu_{i,t} = 1$ (resp. 0), we will refer to opportunity i and volunteer t as compatible (resp. incompatible).

Fixing an EFET β and a parameter α_4, the arrival sequence begins with $\frac{\beta-\alpha_4}{1-\alpha_4}C$ external traffic volunteers for each opportunity $i \in \{1, \ldots, (1-\alpha_4)N\}$ (who are compatible with their targeted opportunity). In total, this comprises $(\beta - \alpha_4)NC$ external traffic volunteers. Next, the internal traffic arrives, which consists of $(1-\beta)NC$ volunteers. These volunteers can be separated into $(1-\alpha_4)N$ batches of size $\frac{1-\beta}{1-\alpha_4}C$ sequentially-arriving homogeneous volunteers, such that the volunteers in the i^{th} batch are compatible with all opportunities $j \geq i$. Finally, additional external traffic arrives, with C compatible volunteers for each opportunity $i \in \{(1-\alpha_4)N + 1, \ldots, N\}$.

We first note that the EFET in such instances is equal to β. To see that this is indeed the case, note that the opportunities in the first subset (those that initially receive external traffic) receive a total filled capacity of $(\beta - \alpha_4)NC$, while the opportunities in the other subset (those that receive external traffic at the end of the arrival sequence) receive a total filled capacity of $\alpha_4 NC$. In sum, this represents a fraction β of total capacity. We now proceed to assessing the value of `MSVV` and `OPT` on this family of instances.

[49] We assume that $\alpha_4 NC$ is an integer. This assumption does not impact the upper bound in the statement of Claim 6, as the expression comes from taking the limit as N approaches ∞.

Value of `MSVV` on Instance $\mathcal{I}_4(\beta)$: We now analyze the value of `MSVV` on this instance. All the initial external traffic will be allocated to the appropriate opportunity. At the conclusion of this process, each opportunity $i \in \{1,\dots,(1-\alpha_4)N\}$ will have a filled capacity of $\frac{\beta-\alpha_4}{1-\alpha_4}C$. Based on the allocation rule of `MSVV`, at first all the internal traffic will be exclusively allocated (evenly) across the other α_4 compatible opportunities, i.e., opportunities $i \in \{(1-\alpha_4)N+1,\dots,N\}$, since those opportunities will have less filled capacity. (Recall that `MSVV` defines an opportunity's fill rate as the ratio of filled capacity to total capacity, regardless of the source of the volunteers.) If there is enough internal traffic to fill all opportunities to an equal fill rate of $\frac{\beta-\alpha_4}{1-\alpha_4}$, then the remaining internal traffic will be evenly split among compatible opportunities, until the internal traffic runs out or the remaining compatible opportunities have all reached capacity. Finally, the external traffic fills opportunities $i \in \{(1-\alpha_4)N+1,\dots,N\}$ to capacity.

This allocation corresponds to two different cases, based on the amount of internal traffic relative to the parameter α_3. To help define these cases, we introduce $\alpha_5 := \min\{1-\alpha_4, \frac{\alpha_4(\beta-\alpha_4)}{1-\beta}\}$. As we will later show, $\alpha_5 N$ represents the highest-indexed opportunity that does not receive internal traffic under `MSVV`. In each case, we will demonstrate that the total amount of filled capacity is given by

$$\sum_{i=1}^{\alpha_5 N}\frac{\beta-\alpha_4}{1-\alpha_4}C+\sum_{i=\alpha_5 N+1}^{(1-\alpha_4)N}\min\left\{\left(\frac{\beta-\alpha_4}{1-\alpha_4}+\sum_{j=\alpha_5 N+1}^{i}\frac{1-\beta}{1-\alpha_4}\cdot\frac{C}{N-j+1}\right),C\right\}+\sum_{i=(1-\alpha_4)N+1}^{N}C.$$

In case (i), the amount of internal traffic is insufficient to equalize the fill rate of all opportunities, i.e., $(1-\beta)C \le \alpha_4\left(\frac{\beta-\alpha_4}{1-\alpha_4}C\right)$. Consequently, `MSVV` will simply divide all internal traffic equally among opportunities $i \in \{(1-\alpha_4)N+1,\dots,N\}$. These opportunities will then be filled to capacity by external traffic. We note that $\alpha_5 = 1-\alpha_4$, since in this case, $1-\alpha_4$ cannot exceed $\frac{\alpha_4(\beta-\alpha_4)}{1-\beta}$. Therefore, the total filled capacity in this case is given by

$$\sum_{i=1}^{\alpha_5 N}\frac{\beta-\alpha_4}{1-\alpha_4}C+\sum_{i=\alpha_5 N+1}^{(1-\alpha_4)N}\min\left\{\left(\frac{\beta-\alpha_4}{1-\alpha_4}+\sum_{j=\alpha_5 N+1}^{i}\frac{1-\beta}{1-\alpha_4}\cdot\frac{C}{N-j+1}\right),C\right\}+\sum_{i=(1-\alpha_4)N+1}^{N}C,$$

as desired. We note that in this case, the middle sum is empty, as $\alpha_5 = 1-\alpha_4$.

In case (ii), the amount of internal traffic is sufficient to equalize all fill rates, i.e., if $(1-\beta)C > \alpha_4\left(\frac{\beta-\alpha_4}{1-\alpha_4}C\right)$ (which implies $\alpha_5 = \frac{\alpha_4(\beta-\alpha_4)}{1-\beta}$). In this case, the opportunities will all reach an equal fill rate after the arrival of the $\alpha_5 N^{\text{th}}$ batch of internal traffic. From this point forward, internal traffic will be split among the compatible opportunities, but none of the first $\alpha_5 N$ opportunities are compatible with remaining arrivals. As such, the total filled capacity is given by

$$\sum_{i=1}^{\alpha_5 N}\frac{\beta-\alpha_4}{1-\alpha_4}C+\sum_{i=\alpha_5 N+1}^{(1-\alpha_4)N}\min\left\{\left(\frac{\beta-\alpha_4}{1-\alpha_4}+\sum_{j=\alpha_5 N+1}^{i}\frac{1-\beta}{1-\alpha_4}\cdot\frac{C}{N-j+1}\right),C\right\}+\sum_{i=(1-\alpha_4)N+1}^{N}C.$$

We now compute the fraction of total capacity that is filled under `MSVV`. To help in this calculation, we define $\alpha_6 := \min\{1-\alpha_4, 1-(1-\alpha_5)/e\}$, which (asymptotically) represents the fraction of opportunities that are not filled to capacity under `MSVV`.

$$\frac{\texttt{MSVV}(\mathcal{I}_4(\beta))}{NC} = \frac{1}{NC}\left[\sum_{i=1}^{\alpha_5 N}\frac{\beta-\alpha_4}{1-\alpha_4}C+\sum_{i=\alpha_5 N+1}^{(1-\alpha_4)N}\min\left\{\left(\frac{\beta-\alpha_4}{1-\alpha_4}+\sum_{j=\alpha_5 N+1}^{i}\frac{1-\beta}{1-\alpha_4}\cdot\frac{C}{N-j+1}\right),C\right\}\right.$$

$$+\sum_{i=(1-\alpha_4)N+1}^{N} C\Bigg] \tag{54}$$

$$\xrightarrow{N\to\infty}\left(\frac{\beta-\alpha_4}{1-\alpha_4}\right)\alpha_5+\int_{\alpha_5}^{1-\alpha_4}\min\left\{\frac{\beta-\alpha_4}{1-\alpha_4}+\frac{1-\beta}{1-\alpha_4}\log\left(\frac{1-\alpha_5}{1-x}\right),1\right\}\,\partial x+\alpha_4 \tag{55}$$

$$=\left(\frac{\beta-\alpha_4}{1-\alpha_4}\right)\alpha_5+\int_{\alpha_5}^{\alpha_6}\frac{\beta-\alpha_4}{1-\alpha_4}+\frac{1-\beta}{1-\alpha_4}\log\left(\frac{1-\alpha_5}{1-x}\right)\,\partial x+\int_{\alpha_6}^{1-\alpha_4}1\,\partial x+\alpha_4 \tag{56}$$

$$=\left(\frac{\beta-\alpha_4}{1-\alpha_4}\right)\alpha_6+\frac{1-\beta}{1-\alpha_4}\int_{\alpha_5}^{\alpha_6}\log\left(\frac{1-\alpha_5}{1-x}\right)\,\partial x+(1-\alpha_6) \tag{57}$$

$$=\frac{\beta-\alpha_4}{1-\alpha_4}\alpha_6+\frac{1-\beta}{1-\alpha_4}\left(\alpha_6-\alpha_5-(1-\alpha_6)\log\left(\frac{1-\alpha_5}{1-\alpha_6}\right)\right)+(1-\alpha_6) \tag{58}$$

$$=1-\frac{1-\beta}{1-\alpha_4}\left(\alpha_5+(1-\alpha_6)\log\left(\frac{1-\alpha_5}{1-\alpha_6}\right)\right) \tag{59}$$

Equality (55) uses the fact that $\sum_{k=yn}^{xn} 1/k$ approaches $\log(x/y)$ as n approaches infinity. Equality (58) comes from applying the definition of α_6 and noting that for $x\geq\alpha_6$, $1\leq\frac{\beta-\alpha_4}{1-\alpha_4}+\frac{1-\beta}{1-\alpha_4}\log\left(\frac{1-\alpha_5}{1-x}\right)$. Taking the integrals and simplifying, we arrive at the final expression, which represents the fraction of total capacity that is filled under `MSVV`.

Value of `OPT` on Instance $\mathcal{I}_4(\beta)$: We now show that `OPT` fills all capacity on this instance. Consider a solution that allocates all external traffic to its targeted opportunity, and allocates the i^{th} batch of internal traffic to opportunity i. Under this solution, each opportunity $i\in\{1,\ldots,(1-\alpha_4)N\}$ receives $\frac{\beta-\alpha_4}{1-\alpha_4}C$ matches from external traffic and $\frac{1-\beta}{1-\alpha_4}C$ matches from internal traffic, thereby reaching its capacity of C. Furthermore, each opportunity $i\in\{(1-\alpha_4)N+1,\ldots,N\}$ receives C matches from external traffic. Thus, under this solution, all capacity is filled, which means that `OPT` must also fill all capacity on this instance.

This establishes a competitive ratio of $1-\frac{1-\beta}{1-\alpha_4}\left(\alpha_5+(1-\alpha_6)\log\left(\frac{1-\alpha_5}{1-\alpha_6}\right)\right)$, as desired. Taking the minimum over all $\alpha_4\in[0,\beta]$ completes the proof of the claim. □

Both claims establish an upper bound on the competitive ratio of `MSVV`.[50] In Figure 2b of Section 4.2, we illustrate the piecewise-defined upper bound on `MSVV` that results from taking the minimum for any particular EFET $\beta>1/e$, along with the universal upper bound of $1-1/e$ for $\beta\leq 1/e$. □

A.6. Omitted Details in the Proof of Theorem 2 (Section 4.4)

A.6.1. Proof of Lemma 3 The proof of Lemma 3 follows from the definition of the pseudo-rewards L_t and K_i (which we replicate below for ease of reference) as well as the definition of the `AC` algorithm.

$$L_t=\begin{cases}\sum_{i\in[n]}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}})=i], & t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0\\ \sum_{i\in[n]}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\xi_t(S_t^{\texttt{OPT}})=i], & t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0\end{cases}$$

$$K_i=\sum_{t\in[T]}(1-\psi(\mathrm{FR}_{i,t-1}))\,\mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}})=i]$$

Recall that $\tilde{\xi}_t(S_t^{\texttt{AC}})$ represents the opportunity that volunteer t *contributes to* under `AC`. To be precise, if opportunity $\xi_t(S_t^{\texttt{AC}})$ has remaining capacity at time t, then $\tilde{\xi}_t(S_t^{\texttt{AC}})=\xi_t(S_t^{\texttt{AC}})$. Otherwise, $\tilde{\xi}_t(S_t^{\texttt{AC}})=0$. In addition, recall that $\mathcal{V}^0$ represents the set of arriving internal traffic for which `OPT` recommends opportunity 0.

[50] To show that these upper bounds hold for any minimum capacity $\underline{c}$, it suffices to add an additional opportunity with capacity $\underline{c}$ for which volunteers have conversion probability of 0. The performance of both `OPT` and `MSVV` are unchanged, and the EFET remains the same in the limit as N approaches infinity.

Based on these definitions,

$$\mathbb{E}_{\boldsymbol{\omega}}[\mathtt{AC}] = \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\sum_{i\in[n]}\mathbb{1}[\tilde{\xi}_t(S_t^{\mathtt{AC}})=i] + \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}\sum_{i\in[n]}\mathbb{1}[\tilde{\xi}_t(S_t^{\mathtt{AC}})=i]\right] \tag{60}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]}\left(\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\tilde{\xi}_t(S_t^{\mathtt{AC}})=i] + \sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}(1-\psi(\mathrm{FR}_{i,t-1}))\mathbb{1}[\tilde{\xi}_t(S_t^{\mathtt{AC}})=i]\right.\right.$$

$$\left.\left. + \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\tilde{\xi}_t(S_t^{\mathtt{AC}})=i] + \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}(1-\psi(\mathrm{FR}_{i,t-1}))\mathbb{1}[\tilde{\xi}_t(S_t^{\mathtt{AC}})=i]\right)\right] \tag{61}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\tilde{\xi}_t(S_t^{\mathtt{AC}})=i]\right] + \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}L_t + \sum_{i\in[n]}K_i\right] \tag{62}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\xi_t(S_t^{\mathtt{AC}})=i]\right] + \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}L_t + \sum_{i\in[n]}K_i\right] \tag{63}$$

$$\geq \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\xi_t(S_t^{\mathtt{OPT}})=i]\right] + \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}L_t + \sum_{i\in[n]}K_i\right] \tag{64}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in[T]}L_t + \sum_{i\in[n]}K_i\right] \tag{65}$$

All steps are algebraic except for (63) and (64). To establish the former, we will show that $\sum_{i\in[n]}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\xi_t(S_t^{\mathtt{AC}}) = i] = \sum_{i\in[n]}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\tilde{\xi}_t(S_t^{\mathtt{AC}}) = i]$. We consider two cases. First, if $\mathrm{FR}_{\xi_t(S_t^{\mathtt{AC}}),t-1} < 1$, then $\xi_t(S_t^{\mathtt{AC}}) = \tilde{\xi}_t(S_t^{\mathtt{AC}})$ and the equality holds. Alternatively, if $\mathrm{FR}_{\xi_t(S_t^{\mathtt{AC}}),t-1} = 1$, then $\tilde{\xi}_t(S_t^{\mathtt{AC}}) = 0$ and $\psi(\mathrm{FR}_{\xi_t(S_t^{\mathtt{AC}}),t-1}) = 0$. Thus, both summations equal 0, and the equality holds.

Inequality (64) follows from the `AC` algorithm's optimality condition (see Algorithm 2), which ensures that it recommends the opportunity that maximizes the weighted probability of generating a sign-up (where the weight for opportunity i at time t is given by $\psi(\mathrm{FR}_{i,t-1})$). Since the recommendation provided by `OPT` to any volunteer must be independent of their sign-up realization, the inequality holds. Applying the definition of the pseudo-rewards L_t for $t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0$ completes the proof of Lemma 3.

A.6.2. Proof of Lemma 4 We will prove a stronger version of this lemma by establishing the following inequality along any fixed sample path $\boldsymbol{\omega}$:

$$\sum_{t\in[T]}L_t + \sum_{i\in[n]}K_i \quad\geq\quad e^{-1/c}\sum_{i\in[n]}\left(\mathtt{AC}_{i,T}^{\text{EXT}} + \mathtt{AC}_{i,T}^{0} + \mathtt{OPT}_{i,T}^{\text{INT}}\cdot\psi\left(\frac{\mathtt{AC}_{i,T}^{\text{INT}}}{c_i - \mathtt{AC}_{i,T}^{\text{EXT}}}\right)\right.$$
$$\left. + c_i\left(1-\psi\left(\frac{\mathtt{AC}_{i,T}^{\text{INT}} - \mathtt{AC}_{i,T}^{0}}{c_i}\right) - 1/e\right)\right),$$

We proceed by separately deriving lower bounds on the L_t pseudo-rewards and the K_i pseudo-rewards. For the former,

$$\sum_{t\in[T]}L_t \quad=\quad \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}L_t + \sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}L_t \tag{66}$$

$$= \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}L_t + \sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\sum_{i\in[n]}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}[\xi_t(S_t^{\mathtt{OPT}})=i] \tag{67}$$

$$\geq \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0} L_t + \sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\sum_{i\in[n]} \psi(\text{FR}_{i,T})\mathbb{1}[\xi_t(S_t^{\texttt{OPT}})=i] \tag{68}$$

$$= \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0} L_t + \sum_{i\in[n]} \psi\left(\frac{\texttt{AC}_{i,T}^{\text{INT}}}{c_i - \texttt{AC}_{i,T}^{\text{EXT}}}\right)\texttt{OPT}_{i,T}^{\text{INT}} \tag{69}$$

Equality in (67) follows from the definition of L_t. Inequality in (68) holds because ψ is a decreasing function in its argument, and $\text{FR}_{i,T}\geq \text{FR}_{i,t-1}$ for all $t\in[T]$. Equality in (69) comes from applying the definition of the fill rate as well as the fact that $\texttt{OPT}_{i,T}^{\text{INT}} = \sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\mathbb{1}[\xi_t(S_t^{\texttt{OPT}})=i]$.

We next turn our attention to the K_i pseudo-rewards, which we further separate into two summations:

$$\sum_{i\in[n]} K_i = \sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0} (1-\psi(\text{FR}_{i,t-1}))\,\mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}})=i] + \sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}(1-\psi(\text{FR}_{i,t-1}))\,\mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}})=i] \tag{70}$$

We note that the first summation has a nice relationship with the first term in (69). To see this, recall that we define $\texttt{AC}_{i,T}^0 = \sum_{t\in\mathcal{V}^0}\mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}})=i]$ as the sum of sign-ups under $\texttt{AC}$ for opportunity i by volunteers who did not receive a recommendation under $\texttt{OPT}$. Then,

$$\sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}(1-\psi(\text{FR}_{i,t-1}))\,\mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}})=i] = \sum_{i\in[n]}\left(\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}\mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}})=i]-\psi(\text{FR}_{i,t-1})\mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}})=i]\right) \tag{71}$$

$$= \sum_{i\in[n]}\texttt{AC}_{i,T}^{\text{EXT}}+\texttt{AC}_{i,T}^0 - \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0} L_t \tag{72}$$

Now focusing on the second summation, which deals with internal traffic for which $\texttt{OPT}$ provides a recommendation:

$$\sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}(1-\psi(\text{FR}_{i,t-1}))\,\mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}})=i] \geq \sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\left(1-\psi\left(\frac{\texttt{AC}_{i,t-1}^{\text{INT}}}{c_i}\right)\right)\mathbb{1}[\tilde{\xi}_t(S_t^{\texttt{AC}})=i] \tag{73}$$

$$\geq \sum_{i\in[n]}\sum_{k\in[\texttt{AC}_{i,T}^{\text{INT}}-\texttt{AC}_{i,T}^0]}\left(1-\psi\left(\frac{k-1}{c_i}\right)\right) \tag{74}$$

$$\geq \sum_{i\in[n]} e^{-1/c_i}\sum_{k\in[\texttt{AC}_{i,T}^{\text{INT}}-\texttt{AC}_{i,T}^0]}\left(1-\psi\left(\frac{k}{c_i}\right)\right) \tag{75}$$

$$\geq e^{-1/\underline{c}}\sum_{i\in[n]}\int_0^{\texttt{AC}_{i,T}^{\text{INT}}-\texttt{AC}_{i,T}^0} 1-\psi(x/c_i)\,\partial x \tag{76}$$

$$= e^{-1/\underline{c}}\sum_{i\in[n]} c_i\left(1-\psi\left(\frac{\texttt{AC}_{i,T}^{\text{INT}}-\texttt{AC}_{i,T}^0}{c_i}\right)-1/e\right) \tag{77}$$

In (73), we use the fact that ψ is decreasing and $\frac{\texttt{AC}_{i,t-1}^{\text{INT}}}{c_i}\leq \frac{\texttt{AC}_{i,t-1}^{\text{INT}}}{c_i-\texttt{AC}_{i,t-1}^{\text{EXT}}} = \text{FR}_{i,t-1}$. We then further reduce the argument in ψ in (74) by noting that the lowest possible values of $\texttt{AC}_{i,t}^{\text{INT}}$ are $\{1,\ldots,\texttt{AC}_{i,T}^{\text{INT}}-\texttt{AC}_{i,T}^0\}$, since $\texttt{AC}_{i,t}^{\text{INT}}$ increases by 1 for any $t\in\mathcal{V}^{\text{INT}}$ where $\tilde{\xi}_t(S_t^{\texttt{AC}})=i$.

The summation in (74) represents a left Reimann sum of an increasing function. In (75), we utilize the fact that for any k, $1-\psi((k-1)/c_i)\geq e^{1/\underline{c}}(1-\psi(k/c_i))$. As the summation in (75) is now a right Reimann sum of an increasing function, we bound the sum with an appropriate integral in (76). We evaluate the integral to arrive at (77).

Combining (69), (72), and (77) along with the observation that $e^{-1/\underline{c}} < 1$, we see that for any sample path $\boldsymbol{\omega}$,

$$\sum_{t\in[T]} L_t + \sum_{i\in[n]} K_i \quad \geq \quad e^{-1/\underline{c}} \sum_{i\in[n]} \Bigg(\mathtt{AC}^{\text{EXT}}_{i,T} + \mathtt{AC}^{0}_{i,T} + \mathtt{OPT}^{\text{INT}}_{i,T}\psi\left(\frac{\mathtt{AC}^{\text{INT}}_{i,T}}{c_i - \mathtt{AC}^{\text{EXT}}_{i,T}}\right) + c_i\left(1 - \psi\left(\frac{\mathtt{AC}^{\text{INT}}_{i,T} - \mathtt{AC}^{0}_{i,T}}{c_i}\right) - 1/e\right)\Bigg)$$

Taking expectations over all sample paths completes the proof of Lemma 4 $\square$

A.6.3. Proof of Lemma 5: To prove Lemma 5, it is sufficient to show that for any instance $\mathcal{I}$, we can construct a feasible solution to (MP) which has a value of $\frac{\mathbb{E}_{\boldsymbol{\omega}}[\mathtt{AC}]}{\mathbb{E}_{\boldsymbol{\omega}}[\mathtt{OPT}]}$. (We remind that AC and OPT depend on both the instance $\mathcal{I}$ and the sample path $\boldsymbol{\omega}$, but we suppress that dependence to ease exposition). For ease of reference, we reproduce (MP) below.

Given an instance $\mathcal{I}$, the inputs to (MP) are the set of opportunities $\mathcal{S}$ and the set of feasible sample paths Ω, along with its associated probability measure.

(MP) uses the set of variables $\vec{x} \in \mathbb{R}^{3\times n\times|\Omega|}_{\geq 0}$ and $\vec{y} \in \mathbb{R}^{2\times n\times|\Omega|}_{\geq 0} \setminus \vec{\mathbf{0}}$, along with $z \in [0,1]$

$$\min_{\vec{x},\vec{y},z} \quad z \qquad \textbf{(MP)}$$

$$\text{s.t.} \quad \forall i,\boldsymbol{\omega}, \quad c_i \geq y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}} \quad \textbf{(i)} \qquad c_i \geq x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}} \quad \textbf{(ii)} \qquad x_{2,i,\boldsymbol{\omega}} \geq x_{3,i,\boldsymbol{\omega}} \quad \textbf{(iii)}$$

$$c_i = x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}} \quad \textbf{OR} \quad x_{1,i,\boldsymbol{\omega}} = y_{1,i,\boldsymbol{\omega}} \qquad \textbf{(iv)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}}\right] \leq z\textstyle\sum_{i\in[n]} c_i \qquad \textbf{(v)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}}\right] \geq (\beta - \sigma + z)\textstyle\sum_{i\in[n]} c_i \qquad \textbf{(vi)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\cdot\psi\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - x_{1,i,\boldsymbol{\omega}}}\right) + c_i\left(1 - \psi\left(\frac{x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}}{c_i}\right) - 1/e\right)\right] \leq e^{1/\underline{c}} z \mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\right] \qquad \textbf{(vii)}$$

To construct such a feasible solution, we define the values of $\vec{x}$ based on the value of AC along a particular sample path.[51] Specifically, $x_{1,i,\boldsymbol{\omega}}$ (resp. $x_{2,i,\boldsymbol{\omega}}$) represents the amount of external traffic (resp. internal traffic) that contributes to opportunity i under AC, given by $\mathtt{AC}^{\text{EXT}}_{i,T}$ (resp. $\mathtt{AC}^{\text{INT}}_{i,T}$). The third component, $x_{3,i,\boldsymbol{\omega}}$, accounts for the value of AC on the volunteers for which OPT recommends opportunity 0, which we denote as $\mathtt{AC}^{0}_{i,T} := \sum_{t\in\mathcal{V}^0} \mathbb{1}[\tilde{\xi}_t(S^{\mathtt{AC}}_t) = i]$. In a similar fashion, we define the values of $\vec{y}$ based on the value of OPT along a particular sample path. Specifically, $y_{1,i,\boldsymbol{\omega}}$ (resp. $y_{2,i,\boldsymbol{\omega}}$) represents the amount of external traffic (resp. internal traffic) that contributes to opportunity i under OPT, given by $\mathtt{OPT}^{\text{EXT}}_{i,T}$ (resp. $\mathtt{OPT}^{\text{INT}}_{i,T}$). Finally, we define z as the ratio between the expected value of AC and the expected value of OPT on this instance.

[51] We emphasize that fixing a sample path $\boldsymbol{\omega}$, the entire sequence of opportunity recommendations and volunteer sign-ups are entirely deterministic under both AC and OPT. To see this, note that for any fixed history, the AC algorithm makes a deterministic recommendation, and the volunteer's decision in response to that recommendation is deterministic, conditional on $\boldsymbol{\omega}$. Similarly, OPT makes a deterministic recommendation for any fixed history and fixed inputs. The history as well as inputs (i.e., the instance $\mathcal{I}$ as well as the sign-up decisions of all external traffic) are deterministic for any fixed $\boldsymbol{\omega}$.

To summarize, we consider the following feasible solution:

$$x_{1,i,\boldsymbol{\omega}} = \mathtt{AC}^{\text{EXT}}_{i,T}, \qquad x_{2,i,\boldsymbol{\omega}} = \mathtt{AC}^{\text{INT}}_{i,T}, \qquad x_{3,i,\boldsymbol{\omega}} = \mathtt{AC}^{0}_{i,T},$$
$$y_{1,i,\boldsymbol{\omega}} = \mathtt{OPT}^{\text{EXT}}_{i,T}, \qquad y_{2,i,\boldsymbol{\omega}} = \mathtt{OPT}^{\text{INT}}_{i,T}, \qquad z = \frac{\mathbb{E}_{\boldsymbol{\omega}}[\mathtt{AC}]}{\mathbb{E}_{\boldsymbol{\omega}}[\mathtt{OPT}]}$$

If such a solution is feasible, then the optimal value of (MP) is at most $\frac{\mathbb{E}_{\boldsymbol{\omega}}[\mathtt{AC}]}{\mathbb{E}_{\boldsymbol{\omega}}[\mathtt{OPT}]}$, since the optimal value of (MP) is less than or equal to the value of any feasible solution. We proceed by sequentially showing that each constraint is met under this candidate solution.[52]

First, observe that neither AC nor OPT can exceed the capacity of the opportunity along any sample path $\boldsymbol{\omega}$. Hence, constraints (i) and (ii) are never violated. Similarly, $\mathtt{AC}^{\text{INT}}_{i,T}$ is the sum of sign-ups from internal traffic under AC, while $\mathtt{AC}^{0}_{i,T}$ is the sum of sign-ups from a subset of internal traffic under AC. Thus, constraint (iii) must hold.

For constraint (iv), we first fix an opportunity i. Based on Definition 1, OPT will never use internal traffic to fill capacity that would otherwise be filled by external traffic. As a consequence, OPT uses all external traffic for i (or fills opportunity i with external traffic) along each sample path. In contrast, AC may use internal traffic to fill capacity that could otherwise have been filled by external traffic. In other words, if an opportunity reaches full capacity under AC, then some external traffic may be excessive. Thus, along a fixed sample path, either AC uses the same amount of external traffic as OPT for opportunity i, or opportunity i reaches capacity under AC.[53] These two possibilities give rise to constraint (iv).

Constraint (v) holds based on the definitions of $\vec{x}, \vec{y}$, and z:

$$z \quad = \quad \frac{\mathbb{E}_{\boldsymbol{\omega}}[\mathtt{AC}]}{\mathbb{E}_{\boldsymbol{\omega}}[\mathtt{OPT}]} \quad = \quad \frac{\mathbb{E}_{\boldsymbol{\omega}}[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}}]}{\mathbb{E}_{\boldsymbol{\omega}}[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}]} \quad \geq \quad \frac{\mathbb{E}_{\boldsymbol{\omega}}[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}}]}{\sum_{i\in[n]} c_i}.$$

We now consider constraint (vi), which crucially provides a lower bound on the number of sign-ups generated by AC where OPT either generates a sign-up to the same opportunity or does not generate a sign-up at all. Fixing a sample path and an opportunity, note that the total amount of opportunity i's capacity filled by AC in periods $t \in \mathcal{V}^{\text{INT}} \setminus \mathcal{V}^0$ is given by $x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}$, while the total amount of opportunity i's capacity filled by OPT in periods $t \in \mathcal{V}^{\text{INT}} \setminus \mathcal{V}^0$ is given by $y_{2,i,\boldsymbol{\omega}}$. Furthermore, for all $t \in \mathcal{V}^{\text{INT}} \setminus \mathcal{V}^0$, OPT provides a recommendation, which means it fills a unit of capacity with probability at least $\underline{\mu}_t$, while AC will fill a unit of capacity with probability at most $\bar{\mu}_t$. As a consequence, $x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}} \leq \sigma y_{2,i,\boldsymbol{\omega}}$, or equivalently, $x_{2,i,\boldsymbol{\omega}} \leq \sigma y_{2,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}}$

Based on the constructed values of $\vec{x}, \vec{y}$, and z, as well as the upper bound on $x_{2,i,\boldsymbol{\omega}}$ identified above,

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}}\right] = z \cdot \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\right] - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{2,i,\boldsymbol{\omega}}\right] \tag{78}$$

$$\geq z \cdot \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\right] - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} \sigma \cdot y_{2,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}}\right] \tag{79}$$

[52] We remark that we restrict our attention to instances where $\mathbb{E}_{\boldsymbol{\omega}}[\mathtt{OPT}] > 0$; thus, $\vec{y}$ can be constrained to have at least one strictly positive element.

[53] By our convention for external traffic, AC will always *recommend* the volunteer's targeted opportunity i^*_t. However, if this opportunity has already reached capacity, the sign-up does not *fill* any capacity.

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}}\right] - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]}(1-z)\cdot y_{1,i,\boldsymbol{\omega}} + (\sigma - z)\cdot y_{2,i,\boldsymbol{\omega}}\right] - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{3,i,\boldsymbol{\omega}}\right] \tag{80}$$

$$\geq \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}}\right] - (\sigma - z)\cdot \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\right] - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{3,i,\boldsymbol{\omega}}\right] \tag{81}$$

$$\geq \beta \sum_{i\in[n]} c_i - (\sigma - z)\sum_{i\in[n]} c_i - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{3,i,\boldsymbol{\omega}}\right]. \tag{82}$$

Inequality (81) uses the fact that $\sigma \geq 1$. The final inequality uses the fact that $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}}\right] = \beta \sum_{i\in[n]} c_i$ based on the definitions of the optimal clairvoyant algorithm $\texttt{OPT}$ and the EFET β (see Definitions 1 and 2). This final inequality establishes that our proposed solution respects constraint (vi).

Finally, we turn our attention to constraint (vii). Given the constructed values of $\vec{x}, \vec{y}$, and z,

$$e^{1/\underline{c}} z \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\right] = e^{1/\underline{c}} \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}}\right] \tag{83}$$

$$\geq e^{1/\underline{c}} \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in[T]} L_t + \sum_{i\in[n]} K_i\right] \tag{84}$$

$$\geq \mathbb{E}_{\boldsymbol{\omega}}\Bigg[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\psi\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - x_{1,i,\boldsymbol{\omega}}}\right)$$

$$+ c_i\left(1 - \psi\left(\frac{x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}}{c_i}\right) - 1/e\right)\Bigg] \tag{85}$$

Inequality (84) comes from applying Lemma 3, while (85) comes from applying Lemma 4. This establishes that constraint (vii) is met under our proposed solution.

In aggregate, we have shown that the proposed solution of $\vec{x}, \vec{y}$, and z are feasible in (MP). This solution attains a value of $z = \frac{\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC}(\mathcal{I},\boldsymbol{\omega})]}{\mathbb{E}_{\boldsymbol{\omega}}[\texttt{OPT}(\mathcal{I},\boldsymbol{\omega})]}$, which completes the proof of Lemma 5. $\square$

A.6.4. Proof of Lemma 6 To prove Lemma 6, we will derive a valid lower bound on the value of (MP) (for a fixed instance $\mathcal{I}$) that is parameterized by the EFET β, the minimum capacity $\underline{c}$, and the maximum heterogeneity across a volunteer's preferences σ. We then argue that our lower bound is uniform given β, $\underline{c}$, and σ, in that it is valid for any given instance $\mathcal{I}$ with those parameters.

To derive the lower bound on the value of (MP), we propose a series of transformations to the optimization problem that will ultimately result in a solvable program. The solution to that transformed program serves as a lower bound on (MP), and its value can be characterized as a function that depends only on the EFET β, the minimum capacity $\underline{c}$, and the maximum heterogeneity across a volunteer's preferences σ. We divide this process into five (algebraic) steps, each of which results in a new formulation for the optimization problem.

First, in **Step (a)** we show there is no feasible solution for $z < e^{-1/\underline{c}}(1 - 1/e)$. Thus, we create a new program ((MP$_a$)) where we restrict the feasible domain. The value of this new program serves as a lower bound on the value of (MP). In **Step (b)**, we show that in (MP$_a$), it is without loss of generality to consider only feasible solutions where constraint (i) binds for all i and $\boldsymbol{\omega}$ pairs. Based on this, we constuct a new program (MP$_b$) which replaces the inequality in constraint (i) with an equality. In **Step (c)**, we relax (MP$_b$)

by replacing constraints (i), (iv), and (vii) with a unified constraint (viii). We define this new program as (MP_c). In **Step (d)**, we transform (MP_c) by replacing the inequalities in constraints (iii), (v), and (vi) with three equalities, thereby creating the program (MP_d). Finally, in **Step (e)**, we convexify the simplified program from the previous step, to arrive at the (solvable) (MP_e). We highlight that the value of each new program serves as a lower bound on the value of the previous program; i.e., the value of (MP_b) is a lower bound on the value of (MP_a), which is a lower bound on the value of (MP).

Step (a): Suppose for a moment that there is a feasible solution where $z < e^{-1/\underline{c}}(1-1/e)$. We will show a contradiction by demonstrating that if such a solution satisfies constraints (ii) and (iv), it cannot satisfy constraint (vii). We begin by fixing a particular opportunity i and a particular sample path $\boldsymbol{\omega}$. If constraint (iv) holds, there are two cases to consider: either $x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}} = c_i$ or $x_{1,i,\boldsymbol{\omega}} = y_{1,i,\boldsymbol{\omega}}$. In the first case, we have that $\psi\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - x_{1,i,\boldsymbol{\omega}}}\right) = 0$, as $\psi(1) = 0$ by definition. Note that the left hand side of constraint (vii) is a weighted summation over opportunities and sample paths, where the weights depend on the probability of the sample path. Let us consider the term in that summation which corresponds to the fixed opportunity i and the fixed sample path $\boldsymbol{\omega}$. This term is bounded by

$$\begin{aligned} x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}} + c_i(1-\psi\left(\frac{x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}}{c_i}\right) - 1/e) &= x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}} + c_i \exp\left(\frac{x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}}{c_i} - 1\right) - \frac{c_i}{e} && (86)\\ &= x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}} + c_i \exp\left(\frac{-x_{1,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}}{c_i}\right) - \frac{c_i}{e} && (87)\\ &\geq x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}} + c_i\left(1 - \frac{x_{1,i,\boldsymbol{\omega}}}{c_i} - \frac{x_{3,i,\boldsymbol{\omega}}}{c_i}\right) - \frac{c_i}{e} && (88)\\ &\geq (1-1/e)c_i && (89)\\ &\geq (1-1/e)(y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}) && (90) \end{aligned}$$

Inequality (88) comes from the fact that $\exp(-x) \geq 1-x$ for all x, and the remaining steps are algebraic.

We now address the second case, where $x_{1,i\boldsymbol{\omega}} = y_{1,i,\boldsymbol{\omega}}$ for this particular i and $\boldsymbol{\omega}$. Let us again consider the term in the summation on the left hand side of constraint (vii) which corresponds to the fixed opportunity i and the fixed sample path $\boldsymbol{\omega}$. This term is bounded by

$$\begin{aligned} &x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\psi\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - x_{1,i,\boldsymbol{\omega}}}\right) + c_i\left(1 - \psi\left(\frac{x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}}{c_i}\right) - 1/e\right)\\ &= \quad y_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}} - y_{2,i,\boldsymbol{\omega}}\exp\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - y_{1,i,\boldsymbol{\omega}}} - 1\right) + c_i\exp\left(\frac{x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}}{c_i} - 1\right) - c_i/e && (91)\\ &\geq \quad y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}} - y_{2,i,\boldsymbol{\omega}}\exp\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - y_{1,i,\boldsymbol{\omega}}} - 1\right) + c_i\exp\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i} - 1\right) - c_i/e && (92) \end{aligned}$$

The second inequality holds because the expression is increasing in $x_{3,i,\boldsymbol{\omega}}$. Note that the right hand side of (92) is quasi-concave in $x_{2,i,\boldsymbol{\omega}}$. We demonstrate quasi-concavity by first noting that the expression is a continuously differentiable function of $x_{2,i,\boldsymbol{\omega}}$, and then by establishing that this function cannot have a local minimum. To prove the latter, we begin by calculating the derivative of the right hand side (RHS) with respect to $x_{2,i,\boldsymbol{\omega}}$.

$$\frac{\partial}{\partial x_{2,i,\boldsymbol{\omega}}}\mathrm{RHS} = \frac{-y_{2,i,\boldsymbol{\omega}}}{c_i - y_{1,i,\boldsymbol{\omega}}}\exp\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - y_{1,i,\boldsymbol{\omega}}} - 1\right) + \exp\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i} - 1\right),$$

which is equal to 0 only when $\frac{y_{2,i,\boldsymbol{\omega}}}{c_i - y_{1,i,\boldsymbol{\omega}}}\exp\left(x_{2,i,\boldsymbol{\omega}}/(c_i - y_{1,i,\boldsymbol{\omega}}) - 1\right) = \exp(x_{2,i,\boldsymbol{\omega}}/c_i - 1)$. When this first-order condition holds, we see that the second derivative of the right hand side with respect to $x_{2,i,\boldsymbol{\omega}}$ must be strictly negative:

$$\frac{\partial^2}{\partial x_{2,i,\boldsymbol{\omega}}^2}\text{RHS} = \frac{-y_{2,i,\boldsymbol{\omega}}}{(c_i - y_{1,i,\boldsymbol{\omega}})^2}\exp\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - y_{1,i,\boldsymbol{\omega}}} - 1\right) + \frac{1}{c_i}\exp\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i} - 1\right) = \frac{-y_{1,i,\boldsymbol{\omega}}}{c_i(c_i - y_{1,i,\boldsymbol{\omega}})}\exp\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i}\right)$$

Hence, this expression is quasi-concave in $x_{2,i,\boldsymbol{\omega}}$, and as a consequence is minimized at one of the extreme points of $x_{2,i,\boldsymbol{\omega}}$.

The two extreme points for $x_{2,i,\boldsymbol{\omega}}$ are 0 and $c_i - x_{1,i,\boldsymbol{\omega}}$ (based on constraint (ii)). If $x_{2,i,\boldsymbol{\omega}} = 0$, the RHS of (92) is equal to $y_{1,i,\boldsymbol{\omega}} + (1-1/e)y_{2,i,\boldsymbol{\omega}}$. If $x_{2,i,\boldsymbol{\omega}} = c_i - x_{1,i,\boldsymbol{\omega}}$, we have returned to the first case for constraint (iv), where we established a lower bound of $(1-1/e)(y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}})$ in (90).

Therefore, we have shown that for any particular i and $\boldsymbol{\omega}$, if constraints (ii) and (iv) are satisfied,

$$x_{1,i,\boldsymbol{\omega}} + (y_{2,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}})\psi\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - x_{1,i,\boldsymbol{\omega}}}\right) + c_i\left(1 - \psi\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i}\right) - 1/e\right) \;\geq\; (1-1/e)(y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}})$$

Summing this up over all opportunities and taking expectations over all sample paths,[54] we see that constraint (vii) must be violated for any $z < e^{-1/\underline{c}}(1-1/e)$. This completes Step (a).

In the subsequent step, we will work with a modified version of (MP), which we refer to as (MP$_a$) (shown below), that restricts the domain by imposing that $z \geq e^{-1/\underline{c}}(1-1/e)$. Any feasible solution to (MP) remains feasible in (MP$_a$), and thus the value of (MP$_a$) is a valid lower bound on the value of (MP).

Given an instance $\mathcal{I}$, the inputs to (MP$_a$) are the set of opportunities $\mathcal{S}$, the EFET β, the MCPR σ, and the set of feasible sample paths Ω, along with its associated probability measure.
(MP$_a$) uses the set of variables $\vec{x} \in \mathbb{R}_{\geq 0}^{3\times n\times|\Omega|}$ and $\vec{y} \in \mathbb{R}_{\geq 0}^{2\times n\times|\Omega|} \setminus \vec{\mathbf{0}}$, along with $z \in [e^{-1/\underline{c}}(1-1/e), 1]$

$$\min_{\vec{x},\vec{y},z} \quad z \qquad (\mathbf{MP}_a)$$

$$\text{s.t.} \quad \forall i,\boldsymbol{\omega}, \quad c_i \geq y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}} \quad \textbf{(i)} \qquad c_i \geq x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}} \quad \textbf{(ii)} \qquad x_{2,i,\boldsymbol{\omega}} \geq x_{3,i,\boldsymbol{\omega}} \quad \textbf{(iii)}$$

$$c_i = x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}} \quad \textbf{OR} \quad x_{1,i,\boldsymbol{\omega}} = y_{1,i,\boldsymbol{\omega}} \qquad \textbf{(iv)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}}\right] \leq z\sum_{i\in[n]} c_i \qquad \textbf{(v)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}}\right] \geq (\beta - \sigma + z)\sum_{i\in[n]} c_i \qquad \textbf{(vi)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\cdot\psi\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - x_{1,i,\boldsymbol{\omega}}}\right) + c_i\left(1 - \psi\left(\frac{x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}}{c_i}\right) - 1/e\right)\right]$$
$$\leq e^{1/\underline{c}} z\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\right] \qquad \textbf{(vii)}$$

Step (b): In this step, we will show that we can restrict our attention to feasible solutions of (MP$_a$) where constraint (i) is tight for all i and $\boldsymbol{\omega}$ without loss of optimality. Consider any feasible solution $\{\vec{x},\vec{y},z\}$ where constraint (i) is loose for some $i,\boldsymbol{\omega}$ pair. We will construct a new solution $\{\vec{x}',\vec{y}',z'\}$ which is feasible and has

[54] Because we restrict our attention to arrival sequences where $\mathbb{E}[\texttt{OPT}]$ is non-zero, this includes at least one opportunity and sample path for which $y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}} > 0$.

the same objective value. Let $y'_{2,i,\boldsymbol{\omega}} = c_i - y_{1,i,\boldsymbol{\omega}}$. The other decision variables are unchanged: $y'_{1,i,\boldsymbol{\omega}} = y_{1,i,\boldsymbol{\omega}}$, $\vec{x}' = \vec{x}$, and $z' = z$.

The objective value is identical in both solutions, and only constraints (i) and (vii) are impacted by (weakly) increasing $y_{2,i,\boldsymbol{\omega}}$ to $y'_{2,i,\boldsymbol{\omega}}$. Constraint (i) is satisfied by construction, and constraint (vii) remains satisfied because for all $x \in [0,1]$, $\psi(x) \le 1 - 1/e \le e^{1/\underline{c}} z$, where the second inequality holds as a result of the restricted domain on z imposed in (MP$_a$). This completes Step (b).

In the subsequent step, we will work with a modified version of (MP$_a$), which we refer to as (MP$_b$) (shown below), that replaces the inequality in constraint (i) with equality. As demonstrated in this step, the tightening of constraint (i) is without loss of optimality; thus, the value of (MP$_b$) is a valid lower bound on the value of (MP$_a$).

Given an instance $\mathcal{I}$, the inputs to (MP$_b$) are the set of opportunities $\mathcal{S}$, the EFET β, the MCPR σ, and the set of feasible sample paths Ω, along with its associated probability measure.

(MP$_b$) uses the set of variables $\vec{x} \in \mathbb{R}_{\ge 0}^{3\times n\times|\Omega|}$ and $\vec{y} \in \mathbb{R}_{\ge 0}^{2\times n\times|\Omega|} \setminus \vec{\mathbf{0}}$, along with $z \in [e^{-1/\underline{c}}(1-1/e), 1]$

$$\min_{\vec{x},\vec{y},z} \quad z \qquad (\mathbf{MP}_b)$$

$$\text{s.t.} \quad \forall i, \boldsymbol{\omega}, \quad c_i = y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}} \quad \textbf{(i)} \qquad c_i \ge x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}} \quad \textbf{(ii)} \qquad x_{2,i,\boldsymbol{\omega}} \ge x_{3,i,\boldsymbol{\omega}} \quad \textbf{(iii)}$$

$$c_i = x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}} \quad \textbf{OR} \quad x_{1,i,\boldsymbol{\omega}} = y_{1,i,\boldsymbol{\omega}} \qquad \textbf{(iv)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}}\right] \le z \textstyle\sum_{i\in[n]} c_i \qquad \textbf{(v)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}}\right] \ge (\beta - \sigma + z) \textstyle\sum_{i\in[n]} c_i \qquad \textbf{(vi)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}} \cdot \psi\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - x_{1,i,\boldsymbol{\omega}}}\right) + c_i\left(1 - \psi\left(\frac{x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}}{c_i}\right) - 1/e\right)\right]$$
$$\le e^{1/\underline{c}} z \mathbb{E}_{\boldsymbol{\omega}}\left[\textstyle\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\right] \qquad \textbf{(vii)}$$

Step (c): We will show that we can relax (MP$_b$) by replacing constraints (i), (iv), and (vii) with the following constraint:

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} c_i \hat{g}\left(\frac{x_{1,i,\boldsymbol{\omega}}}{c_i}, \frac{x_{2,i,\boldsymbol{\omega}}}{c_i}, \frac{x_{3,i,\boldsymbol{\omega}}}{c_i}\right)\right] \le e^{1/\underline{c}} z \sum_{i\in[n]} c_i, \qquad \textbf{(viii)}$$

where

$$\hat{g}(x_1, x_2, x_3) = x_1 + x_3 + (1 - x_1) \cdot \psi\left(\frac{x_2}{1 - x_1}\right) + 1 - \psi(x_2 - x_3) - 1/e. \qquad (93)$$

This relaxation results in a new program, which we refer to as (MP$_c$). We now prove that the value of (MP$_c$) provides a lower bound on the value of (MP$_b$) by showing that any solution which satisfies constraints (i), (iv), and (vii) must necessarily satisfy constraint (viii). In (MP$_b$), constraint (i) binds, which means that the right hand sides of constraints (vii) and (viii) are identical. The difference between the left hand sides of constraints (vii) and (viii) is simply the expected sum of $(y_{2,i,\boldsymbol{\omega}} - c_i + x_{1,i,\boldsymbol{\omega}}) \cdot \psi\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - x_{1,i,\boldsymbol{\omega}}}\right)$. Given a solution where constraint (iv) is satisfied for every $i, \boldsymbol{\omega}$ pair, we must have either $\psi\left(\frac{x_{2,i,\boldsymbol{\omega}}}{c_i - x_{1,i,\boldsymbol{\omega}}}\right) = 0$, or

$c_i - x_{1,i,\boldsymbol{\omega}} = c_i - y_{1,i,\boldsymbol{\omega}} = y_{2,i,\boldsymbol{\omega}}$. (The second equality comes from the fact that constraint (i) binds.) As a consequence, the difference between the left hand sides of constraints (vii) and (viii) is 0. Thus, any solution satisfying constraints (i), (iv), and (vii) must also satisfy constraint (viii). This completes step (c), and in the subsequent step, we will work with (MP$_c$) (shown below). We note that the variables $\vec{y}$ do not appear in either the objective or the constraints of (MP$_c$). As a result, we remove these variables from the program.

Given an instance $\mathcal{I}$, the inputs to (MP$_c$) are the set of opportunities $\mathcal{S}$, the EFET β, the MCPR σ, and the set of feasible sample paths Ω, along with its associated probability measure.

(MP$_c$) uses the set of variables $\vec{x} \in \mathbb{R}_{\geq 0}^{3\times n\times|\Omega|}$ and $z \in [e^{-1/\underline{c}}(1-1/e), 1]$

$$\min_{\vec{x}, z} \quad z \qquad \textbf{(MP}_c\textbf{)}$$

$$\text{s.t.} \quad \forall i, \boldsymbol{\omega}, \quad c_i \geq x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}} \quad \textbf{(ii)} \qquad x_{2,i,\boldsymbol{\omega}} \geq x_{3,i,\boldsymbol{\omega}} \quad \textbf{(iii)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum\nolimits_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}}\right] \leq z\sum\nolimits_{i\in[n]} c_i \quad \textbf{(v)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum\nolimits_{i\in[n]} x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}}\right] \geq (\beta - \sigma + z)\sum\nolimits_{i\in[n]} c_i \quad \textbf{(vi)}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum\nolimits_{i\in[n]} c_i \hat{g}\left(\frac{x_{1,i,\boldsymbol{\omega}}}{c_i}, \frac{x_{2,i,\boldsymbol{\omega}}}{c_i}, \frac{x_{3,i,\boldsymbol{\omega}}}{c_i}\right)\right] \leq e^{1/\underline{c}} z\sum\nolimits_{i\in[n]} c_i \quad \textbf{(viii)}$$

Step (d): In this step, we transform (MP$_c$) by replacing constraints (iii), (v), and (vi) with equalities for $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}}\right]$, $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{2,i,\boldsymbol{\omega}}\right]$, and $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{3,i,\boldsymbol{\omega}}\right]$. We will show that such a transformation is without loss of optimality, and we will refer to the resulting program as (MP$_d$). To aid in this step, below we compute the derivatives of $\hat{g}(x_1, x_2, x_3)$, as defined in (93).

$$\frac{\partial \hat{g}}{\partial x_1} = \exp\left(\frac{x_2}{1-x_1} - 1\right)\left(1 - \frac{x_2}{1-x_1}\right) \tag{94}$$

$$\frac{\partial \hat{g}}{\partial x_2} = -\exp\left(\frac{x_2}{1-x_1} - 1\right) + \exp\left(x_2 - x_3 - 1\right) \tag{95}$$

$$\frac{\partial \hat{g}}{\partial x_3} = 1 - \exp\left(x_2 - x_3 - 1\right) \tag{96}$$

Based on these derivatives, we can replace constraints (iii), (v), and (vi) with the following constraints:

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}}\right] = \max\{0, \beta - \sigma + z\}\sum_{i\in[n]} c_i \tag{ix}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{2,i,\boldsymbol{\omega}}\right] = (z - \max\{0, \beta - \sigma + z\})\sum_{i\in[n]} c_i \tag{x}$$

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{3,i,\boldsymbol{\omega}}\right] = 0 \tag{xi}$$

To see why, first consider any feasible solution for (MP$_c$), $\{\vec{x}, z\}$, such that $x_{3,i,\boldsymbol{\omega}} > 0$ for some $i, \boldsymbol{\omega}$ pair. We construct a new solution $\{\vec{x}', z'\}$, where $x'_{1,i,\boldsymbol{\omega}} = x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}}$, $x'_{2,i,\boldsymbol{\omega}} = x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}$, $x'_{3,i,\boldsymbol{\omega}} = 0$ and all other variables remain the same, including $z' = z$. Clearly, this solution has an equivalent objective value, and we can show that such a solution remains feasible.

For constraint (ii) and constraint (v), note that $x'_{1,i,\boldsymbol{\omega}} + x'_{2,i,\boldsymbol{\omega}} = x_{1,i,\boldsymbol{\omega}} + x_{2,i,\boldsymbol{\omega}}$. Similarly, for constraint (iii), note that $x'_{2,i,\boldsymbol{\omega}} - x'_{3,i,\boldsymbol{\omega}} = x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}$, and for constraint (vi), we have $x'_{1,i,\boldsymbol{\omega}} + x'_{3,i,\boldsymbol{\omega}} = x_{1,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}}$. Finally, note that based on the derivatives calculated above, any increase in x_1 and proportional decrease in x_2 and x_3 must (weakly) decrease the left hand side of constraint (viii):

$$\begin{aligned}
\frac{\partial \hat{g}}{\partial x_1} - \frac{\partial \hat{g}}{\partial x_2} - \frac{\partial \hat{g}}{\partial x_3} &= \exp\left(\frac{x_2}{1-x_1} - 1\right)\left(2 - \frac{x_2}{1-x_1}\right) - 1 && (97)\\
&= \exp\left(\frac{x_2}{1-x_1} - 1\right)\left(2 - \frac{x_2}{1-x_1} - \exp\left(1 - \frac{x_2}{1-x_1}\right)\right) && (98)\\
&\leq \exp\left(\frac{x_2}{1-x_1} - 1\right)\left(2 - \frac{x_2}{1-x_1} - 2 + \frac{x_2}{1-x_1}\right) && (99)\\
&\leq 0 && (100)
\end{aligned}$$

Note that the second-to-last inequality uses the fact that $e^x \geq 1 + x$ for any x. This proves that the constructed solution remains feasible, and thus it is without loss of optimality to impose the constraint that $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{3,i,\boldsymbol{\omega}}\right] = 0$.

Using a similar approach that relies on the fact that $\frac{\partial \hat{g}}{\partial x_1} \geq 0$, we can show that it is without loss of generality to assume that either constraint (vi) binds or (if the right hand side of constraint (vi) is negative) every $x_{1,i,\boldsymbol{\omega}} = 0$. Otherwise, we can simply reduce any non-zero $x_{1,i,\boldsymbol{\omega}}$ and remain feasible. Coupled with the constraint $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{3,i,\boldsymbol{\omega}}\right] = 0$, this establishes the equality for $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}}\right]$.

Again using a similar approach, this time relying on the fact that $\frac{\partial \hat{g}}{\partial x_2} \leq 0$, we can show that it is without loss of generality to assume that constraint (v) binds. Otherwise, we can simply increase any $x_{2,i,\boldsymbol{\omega}}$ where constraint (ii) is loose (such an $i,\boldsymbol{\omega}$ pair must exist if constraint (v) is loose). Coupled with the constraint on $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}}\right]$, this establishes the equality for $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{2,i,\boldsymbol{\omega}}\right]$.

Therefore, we can impose the three equality constraints without loss of optimality, and we can then relax the program by dropping constraints (iii), (v), and (vi). This transforms (MP_c) into a new program (MP_d) (shown below), where the value of (MP_d) is a lower bound on the value of (MP_c). This completes step (d), and for the next and final step, we will use (MP_d) as the starting point.

Given an instance $\mathcal{I}$, the inputs to (MP_d) are the set of opportunities $\mathcal{S}$, the EFET β, the MCPR σ, and the set of feasible sample paths Ω, along with its associated probability measure.

(MP_d) uses the set of variables $\vec{x} \in \mathbb{R}_{\geq 0}^{3\times n\times|\Omega|}$ and $z \in [e^{-1/\underline{c}}(1-1/e), 1]$

$$\begin{aligned}
\min_{\vec{x},z} \quad & z && (\mathbf{MP}_d)\\
\text{s.t.} \quad & \forall i,\boldsymbol{\omega}, \quad x_{2,i,\boldsymbol{\omega}} \geq x_{3,i,\boldsymbol{\omega}} && \textbf{(iii)}\\
& \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} c_i \hat{g}\left(\frac{x_{1,i,\boldsymbol{\omega}}}{c_i}, \frac{x_{2,i,\boldsymbol{\omega}}}{c_i}, \frac{x_{3,i,\boldsymbol{\omega}}}{c_i}\right)\right] \leq e^{1/\underline{c}} z \sum_{i\in[n]} c_i && \textbf{(viii)}\\
& \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}}\right] = \max\{0, \beta - \sigma + z\} \sum_{i\in[n]} c_i && \textbf{(ix)}\\
& \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{2,i,\boldsymbol{\omega}}\right] = (z - \max\{0, \beta - \sigma + z\}) \sum_{i\in[n]} c_i && \textbf{(x)}\\
& \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{3,i,\boldsymbol{\omega}}\right] = 0 && \textbf{(xi)}
\end{aligned}$$

Step (e): In the final step, we relax (MP_d) by replacing $\hat{g}(x_1, x_2, 0) := g(x_1, x_2)$ with its lower convex envelope over the domain $\mathcal{D} = \{(x_1, x_2) \in \mathbb{R}^2_{\geq 0} : x_1 + x_2 \leq 1\}$. We denote this lower convex envelope by $\breve{g}(x_1, x_2)$. Any solution which satisfies constraint (viii) in (MP_d) will continue to satisfy constraint (viii) after this change, due to the lower convex envelope being a lower bound (by definition) on the original function g.

Furthermore, as the function $\breve{g}$ is convex, we can require constraints (ix), (x), and (xi) to hold pointwise (i.e., for any $i, \boldsymbol{\omega}$ pair) without loss of optimality. To see why, note that any feasible solution in (MP_d) will remain feasible when averaging over opportunities and sample paths such that $\frac{x_{1,i,\boldsymbol{\omega}}}{c_i}$ is the same for all $i, \boldsymbol{\omega}$ pairs. (This averaging would not impact the value of the solution, z). Similarly, any feasible solution in (MP_d) will remain feasible when averaging over opportunities and sample paths such that $\frac{x_{2,i,\boldsymbol{\omega}}}{c_i}$ is the same for all $i, \boldsymbol{\omega}$ pairs. Additionally, constraint (xi) ensures that $x_{3,i,\boldsymbol{\omega}} = 0$ for all $i, \boldsymbol{\omega}$ pairs, which eliminates the need for constraint (iii).

Based on these observations, we can construct a new program, which we denote by (MP_e), where $x_{1,i,\boldsymbol{\omega}} = \max\{0, \beta - \sigma + z\}$, $x_{2,i,\boldsymbol{\omega}} = z - \max\{0, \beta - \sigma + z\}$, and $x_{3,i,\boldsymbol{\omega}} = 0$ for all $i, \boldsymbol{\omega}$ pairs. We then plug these values into constraint (viii), the only remaining constraint, to arrive at (MP_e) (shown below). As this transformation was without loss of optimality, we note that (MP_e) (shown below) represents a lower bound on (MP_d).

Given an instance $\mathcal{I}$, the inputs to (MP_e) are the EFET β, the minimum capacity $\underline{c}$, and the MCPR σ.

(MP_e) uses the variable $z \in [e^{-1/\underline{c}}(1 - 1/e), 1]$

$$\min_{z} \quad z \qquad\qquad (\mathbf{MP}_e)$$

$$\text{s.t.} \quad \breve{g}\left(\max\{0, \beta - \sigma + z\}, z - \max\{0, \beta - \sigma + z\}\right) \leq e^{1/\underline{c}} z \qquad \textbf{(viii)}$$

We note that the value of (MP_e) is equivalent to z^*, as defined in (6) (see Theorem 2). Furthermore, by steps (a) through (e) and the transitivity property, we have shown that the value of (MP_e) represents a lower bound on the value of (MP) for any instance $\mathcal{I}$. We emphasize that this lower bound depends only on the EFET β of the instance, the minimum capacity $\underline{c}$ of the instance, and the maximum heterogeneity across a volunteer's preferences σ of the instance. This completes the proof of Lemma 6. $\square$

Appendix B: Omitted Proofs of Section 5

B.1. Proof of Proposition 4 (Section 5)

The proof of Proposition 4 follows an identical approach to the proof of Theorem 2. However, it does not require the machinery of Step 3 in the proof of Lemma 2, as we do not intend to break the barrier of $1 - 1/e$ except in trivial cases where the EFET exceeds $1 - 1/e$. Up to that point (i.e., Step 3), this proof follows the exact steps of the proof of Theorem 2. From that point, we complete the proof of Proposition 4 by placing a further lower bound on the value of the `AC-R` algorithm that no longer depends on the amount of capacity filled by external traffic (see Lemma 13).

To begin, we note that even in this ranking setting, if the EFET is β, then the `AC-R` algorithm will fill at least a β fraction of capacity.

Lemma 9 *Let the smallest capacity be given by* $\underline{c}$*. Then, for any effective fraction of external traffic* β*, the competitive ratio of the* AC-R *algorithm is at least* β*.*

Proof of Lemma 9: The proof of Lemma 9 is immediate and is identical to the proof of Lemma 1. We simply note that the AC-R algorithm always recommends the targeted opportunity to external traffic. Applying the definition of the EFET (see Definition 2), this feature of the AC-R algorithm ensures that at least a β fraction of capacity is filled in expectation. $\square$

Next, we establish a lower bound of $e^{-1/\underline{c}}(1-1/e)$ on the competitive ratio of the AC-R algorithm, which requires more intricate analysis.

Lemma 10 *Let the smallest capacity be given by* $\underline{c}$*. Then, for any effective fraction of external traffic* β*, the competitive ratio of the* AC-R *algorithm is at least* $e^{-1/\underline{c}}(1-1/e)$*.*

Proof of Lemma 10: Fixing an instance $\mathcal{I}$, we aim to lower-bound the expected amount of capacity filled under the AC-R algorithm, where the expectation is taken over *sample paths*. In the ranking setting, we extend our definition of a sample path such that $\boldsymbol{\omega} = \{\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_T\}$ represents the realizations of random variables that govern volunteer sign-up decisions. More specifically, we define $\boldsymbol{\omega}_t$ as a vector of length $|\mathcal{S}^{\mathcal{R}}|$ (i.e., $\boldsymbol{\omega}_t$ has one component for every possible ranked set of recommendations). The component of $\boldsymbol{\omega}_t$ corresponding to the ranking $\vec{S} \in \mathcal{S}^{\mathcal{R}}$ indicates the opportunity $i \in \mathcal{S} \cup \{0\}$ that volunteer t signs up for, conditional on the platform recommending the ranked subset $\vec{S}$.[55]

For a fixed instance $\mathcal{I}$ and a fixed sample path $\boldsymbol{\omega}$, we use AC-R to denote the amount of capacity filled under the AC-R algorithm.[56] To provide a lower bound on $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC-R}]$, we leverage the LP-free approach developed in Goyal and Udwani (2019) and Goyal et al. (2020), which involves the creation of path-based pseudo-rewards. (For a more complete discussion of the intuition behind this approach, we kindly refer to the proof sketch of Theorem 2 in Section 4.4.)

Before defining our pseudo-rewards in this setting, recall our convention that any algorithm (including OPT and the AC-R algorithm) always recommends the targeted opportunity to external traffic. As before, to ensure that we do not count sign-ups that exceed the capacity of an opportunity, we define $\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})$ as the opportunity that volunteer t *fills capacity of* under AC-R.

Furthermore, recall that for a fixed instance $\mathcal{I}$ and along a fixed sample path $\boldsymbol{\omega}$, we denote by $\mathcal{V}^0$ the subset of internal traffic for which OPT recommends the dummy ranking $\{0\}$; i.e., OPT does not recommend any opportunity. (Recall that OPT knows *a priori* how much capacity will be filled by external traffic as it knows the realizations of those volunteers' sign-up decisions. This capacity is effectively reserved for external traffic, and internal traffic will be used only if it can fill the remaining capacity. See Definition 1 and its following discussion.)

[55] Fixing a sample path $\boldsymbol{\omega}$, the output of OPT and AC-R are deterministic.

[56] Even though AC-R depends on the instance and the sample path, we hereafter suppress this dependence to ease exposition (for AC-R as well as for all other quantities that depend on the instance and the sample path).

For the fixed instance $\mathcal{I}$ and the fixed sample path $\boldsymbol{\omega}$, we define the pseudo-rewards $L_t^{\mathcal{R}}$ for all $t \in [T]$ and $K_i^{\mathcal{R}}$ for all $i \in [n]$ according to the following:

$$L_t^{\mathcal{R}}(\boldsymbol{\omega}) = \begin{cases} \sum_{i\in[n]} \psi(\mathrm{FR}_{i,t-1}) \mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i], & t \in \mathcal{V}^{\mathrm{EXT}} \cup \mathcal{V}^0 \\ \sum_{i\in[n]} \psi(\mathrm{FR}_{i,t-1}) \mathbb{1}[\xi_t(\vec{S}_t^{\texttt{OPT}}) = i], & t \in \mathcal{V}^{\mathrm{INT}} \setminus \mathcal{V}^0 \end{cases} \tag{101}$$

$$K_i^{\mathcal{R}}(\boldsymbol{\omega}) = \sum_{t\in[T]} (1 - \psi(\mathrm{FR}_{i,t-1}))\, \mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i] \tag{102}$$

We now prove that the expected sum of these pseudo-rewards serves as a lower bound on the expected value of `AC-R`.

Lemma 11 *For any instance $\mathcal{I}$,*

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\texttt{AC-R}\right] \quad \geq \quad \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in[T]} L_t^{\mathcal{R}} + \sum_{i\in[n]} K_i^{\mathcal{R}}\right], \tag{103}$$

where $L_t^{\mathcal{R}}$ and $K_i^{\mathcal{R}}$ are defined in (101) and (102), respectively.

Proof of Lemma 11: The proof follows from the definition of $L_t^{\mathcal{R}}$ and $K_i^{\mathcal{R}}$, as well as the design of the `AC-R` algorithm:

$$\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC-R}] = \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\mathrm{INT}}\setminus\mathcal{V}^0} \sum_{i\in[n]} \mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i] + \sum_{t\in\mathcal{V}^{\mathrm{EXT}}\cup\mathcal{V}^0} \sum_{i\in[n]} \mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i]\right] \tag{104}$$

$$\begin{aligned} &= \mathbb{E}_{\boldsymbol{\omega}}\Bigg[\sum_{i\in[n]} \Bigg(\sum_{t\in\mathcal{V}^{\mathrm{INT}}\setminus\mathcal{V}^0} \psi(\mathrm{FR}_{i,t-1}) \mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i] + \sum_{t\in\mathcal{V}^{\mathrm{INT}}\setminus\mathcal{V}^0} (1-\psi(\mathrm{FR}_{i,t-1}))\mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i] \\ &\qquad + \sum_{t\in\mathcal{V}^{\mathrm{EXT}}\cup\mathcal{V}^0} \psi(\mathrm{FR}_{i,t-1}) \mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i] + \sum_{t\in\mathcal{V}^{\mathrm{EXT}}\cup\mathcal{V}^0} (1-\psi(\mathrm{FR}_{i,t-1}))\mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i]\Bigg)\Bigg] \end{aligned} \tag{105}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} \sum_{t\in\mathcal{V}^{\mathrm{INT}}\setminus\mathcal{V}^0} \psi(\mathrm{FR}_{i,t-1}) \mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i]\right] + \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\mathrm{EXT}}\cup\mathcal{V}^0} L_t^{\mathcal{R}} + \sum_{i\in[n]} K_i^{\mathcal{R}}\right] \tag{106}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} \sum_{t\in\mathcal{V}^{\mathrm{INT}}\setminus\mathcal{V}^0} \psi(\mathrm{FR}_{i,t-1}) \mathbb{1}[\xi_t(\vec{S}_t^{\texttt{AC-R}}) = i]\right] + \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\mathrm{EXT}}\cup\mathcal{V}^0} L_t^{\mathcal{R}} + \sum_{i\in[n]} K_i^{\mathcal{R}}\right] \tag{107}$$

$$\geq \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} \sum_{t\in\mathcal{V}^{\mathrm{INT}}\setminus\mathcal{V}^0} \psi(\mathrm{FR}_{i,t-1}) \mathbb{1}[\xi_t(\vec{S}_t^{\texttt{OPT}}) = i]\right] + \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\mathrm{EXT}}\cup\mathcal{V}^0} L_t^{\mathcal{R}} + \sum_{i\in[n]} K_i^{\mathcal{R}}\right] \tag{108}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in[T]} L_t^{\mathcal{R}} + \sum_{i\in[n]} K_i^{\mathcal{R}}\right] \tag{109}$$

All steps are algebraic except for (107) and Line (108). To establish the former, we will show that $\sum_{i\in[n]} \psi(\mathrm{FR}_{i,t-1}) \mathbb{1}[\xi_t(\vec{S}_t^{\texttt{AC-R}}) = i] = \sum_{i\in[n]} \psi(\mathrm{FR}_{i,t-1}) \mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i]$ for $t \in \mathcal{V}^{\mathrm{INT}} \cup \mathcal{V}^0$. We consider two cases. First, if $\mathrm{FR}_{\xi_t(\vec{S}_t^{\texttt{AC-R}}),t-1} < 1$, then $\xi_t(\vec{S}_t^{\texttt{AC-R}}) = \tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})$ and the equality holds. Alternatively, if $\mathrm{FR}_{\xi_t(\vec{S}_t^{\texttt{AC-R}}),t-1} = 1$, then $\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = 0$ and $\psi(\mathrm{FR}_{\xi_t(\vec{S}_t^{\texttt{AC-R}}),t-1}) = 0$. Thus, both summations equal 0, and the equality holds.

Inequality (108) follows from the `AC-R` algorithm's optimality condition (see Equation 12), which ensures that it recommends the ranking that maximizes the weighted probability of generating a sign-up (where the weight for opportunity i at time t is given by $\psi(\mathrm{FR}_{i,t-1})$). Since the recommendation provided by `OPT` to any volunteer must be independent of their sign-up realization, the inequality holds.[57] Applying the definition of the pseudo-rewards $L_t^{\mathcal{R}}$ for $t \in \mathcal{V}^{\text{INT}} \setminus \mathcal{V}^0$ completes the proof of Lemma 11. □

Next, we place a lower bound on the expected sum of the pseudo-rewards, which depends on the amount of capacity filled under $\texttt{OPT}$ along a fixed sample path. As part of this lower bound, we define $\texttt{AC-R}_{i,t}^{\text{INT}}$ (as well as $\texttt{AC-R}_{i,t}^{\text{EXT}}$ and $\texttt{OPT}_{i,t}^{\text{INT}}$) in exactly the same way as its counterpart in our base model, i.e., as the amount of opportunity i's capacity filled at time t by internal traffic under $\texttt{AC-R}$.

Lemma 12 *For any instance $\mathcal{I}$,*

$$\mathbb{E}_\omega\left[\sum_{t\in[T]} L_t^{\mathcal{R}} + \sum_{i\in[n]} K_i^{\mathcal{R}}\right] \geq e^{-1/c}\mathbb{E}_\omega\left[\sum_{i\in[n]} \texttt{AC-R}_{i,T}^{\text{EXT}} + \texttt{AC-R}_{i,T}^{0} + \texttt{OPT}_{i,T}^{\text{INT}}\cdot\psi\left(\frac{\texttt{AC-R}_{i,T}^{\text{INT}}}{c_i - \texttt{AC-R}_{i,T}^{\text{EXT}}}\right)\right.$$
$$\left. + c_i\left(1-\psi\left(\frac{\texttt{AC-R}_{i,T}^{\text{INT}} - \texttt{AC-R}_{i,T}^{0}}{c_i}\right) - 1/e\right)\right], \tag{110}$$

where $L_t^{\mathcal{R}}$ and $K_i^{\mathcal{R}}$ are defined in (101) and (102), respectively.

Proof of Lemma 12: We proceed by separately deriving lower bounds on the $L_t^{\mathcal{R}}$ pseudo-rewards and the $K_i^{\mathcal{R}}$ pseudo-rewards. For the former,

$$\begin{aligned}
\sum_{t\in[T]} L_t^{\mathcal{R}} &= \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0} L_t^{\mathcal{R}} + \sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0} L_t^{\mathcal{R}} & (111)\\
&= \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0} L_t^{\mathcal{R}} + \sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\sum_{i\in[n]} \psi(\text{FR}_{i,t-1})\mathbb{1}[\xi_t(\vec{S}_t^{\texttt{OPT}}) = i] & (112)\\
&\geq \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0} L_t^{\mathcal{R}} + \sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\sum_{i\in[n]} \psi(\text{FR}_{i,T})\mathbb{1}[\xi_t(\vec{S}_t^{\texttt{OPT}}) = i] & (113)\\
&= \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0} L_t^{\mathcal{R}} + \sum_{i\in[n]} \psi\left(\frac{\texttt{AC-R}_{i,T}^{\text{INT}}}{c_i - \texttt{AC-R}_{i,T}^{\text{EXT}}}\right)\texttt{OPT}_{i,T}^{\text{INT}} & (114)
\end{aligned}$$

Equality in (112) follows from the definition of $L_t^{\mathcal{R}}$. Inequality in (113) holds because ψ is a decreasing function in its argument and $\text{FR}_{i,T} \geq \text{FR}_{i,t-1}$ for all $t\in[T]$. Equality in (114) comes from applying the definition of the fill rate as well as the fact that $\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\mathbb{1}[\xi_t(\vec{S}_t^{\texttt{OPT}}) = i] = \texttt{OPT}_{i,T}^{\text{INT}}$.

We next turn our attention to the $K_i^{\mathcal{R}}$ pseudo-rewards, which we further separate into two summations:

$$\sum_{i\in[n]} K_i^{\mathcal{R}} = \sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}(1-\psi(\text{FR}_{i,t-1}))\,\mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i] + \sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}(1-\psi(\text{FR}_{i,t-1}))\,\mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i] \tag{115}$$

We note that the first summation has a nice relationship with the first term in (114). To see this, let us define $\texttt{AC-R}_{i,T}^0 = \sum_{t\in\mathcal{V}^0}\mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i]$ as the sum of sign-ups under $\texttt{AC-R}$ by volunteers who did not receive a ranking under $\texttt{OPT}$. Then,

$$\sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}(1-\psi(\text{FR}_{i,t-1}))\,\mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i] = \sum_{i\in[n]}\left(\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0}\mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i] - \psi(\text{FR}_{i,t-1})\mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i]\right) \tag{116}$$
$$= \sum_{i\in[n]}\texttt{AC-R}_{i,T}^{\text{EXT}} + \texttt{AC-R}_{i,T}^0 - \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^0} L_t^{\mathcal{R}} \tag{117}$$

[57] We emphasize that, similar to our base model, $\texttt{OPT}$ has knowledge of the arrival sequence and the realized decisions of external traffic, but not the realized decisions of internal traffic.

Now focusing on the second summation, which deals with internal traffic for which $\texttt{OPT}$ provides a ranking:

$$\begin{aligned}
\sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}(1-\psi(\text{FR}_{i,t-1}))\,\mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})=i] &\geq \sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^0}\left(1-\psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,t-1}}{c_i}\right)\right)\mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})=i] &(118)\\
&\geq \sum_{i\in[n]}\sum_{k\in[\texttt{AC-R}^{\text{INT}}_{i,T}-\texttt{AC-R}^0_{i,T}]}\left(1-\psi\left(\frac{k-1}{c_i}\right)\right) &(119)\\
&\geq \sum_{i\in[n]}e^{-1/c_i}\sum_{k\in[\texttt{AC-R}^{\text{INT}}_{i,T}-\texttt{AC-R}^0_{i,T}]}\left(1-\psi\left(\frac{k}{c_i}\right)\right) &(120)\\
&\geq e^{-1/\underline{c}}\sum_{i\in[n]}\int_0^{\texttt{AC-R}^{\text{INT}}_{i,T}-\texttt{AC-R}^0_{i,T}}1-\psi(x/c_i)\,\partial x &(121)\\
&= e^{-1/\underline{c}}\sum_{i\in[n]}c_i\left(1-\psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}-\texttt{AC-R}^0_{i,T}}{c_i}\right)-1/e\right) &(122)
\end{aligned}$$

In (118), we use the fact that ψ is decreasing and $\frac{\texttt{AC-R}^{\text{INT}}_{i,t-1}}{c_i}\leq\frac{\texttt{AC-R}^{\text{INT}}_{i,t-1}}{c_i-\texttt{AC-R}^{\text{EXT}}_{i,t-1}}=\text{FR}_{i,t-1}$. We then further reduce the argument in ψ in (119) by noting that the lowest possible values of $\texttt{AC-R}^{\text{INT}}_{i,t}$ are $\{1,\dots,\texttt{AC-R}^{\text{INT}}_{i,T}-\texttt{AC-R}^0_{i,T}\}$, since $\texttt{AC-R}^{\text{INT}}_{i,t}$ increases by 1 for any $t\in\mathcal{V}^{\text{INT}}$ where $\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})=i$.

The summation in (119) represents a left Reimann sum of an increasing function. In (120), we utilize the fact that for any k, $1-\psi((k-1)/c_i)\geq e^{1/\underline{c}}(1-\psi(k/c_i))$. As the summation in (120) is now a right Reimann sum of an increasing function, we bound the sum with an appropriate integral in (121). We evaluate the integral to arrive at (122).

Combining (114), (117), and (122) along with the observation that $e^{-1/\underline{c}}<1$, we see that for any sample path $\boldsymbol{\omega}$,

$$\begin{aligned}
\sum_{t\in[T]}L_t^{\mathcal{R}}+\sum_{i\in[n]}K_i^{\mathcal{R}} \quad\geq\quad e^{-1/\underline{c}}\sum_{i\in[n]}\Bigg(&\texttt{AC-R}^{\text{EXT}}_{i,T}+\texttt{AC-R}^0_{i,T}+\texttt{OPT}^{\text{INT}}_{i,T}\cdot\psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i-\texttt{AC-R}^{\text{EXT}}_{i,T}}\right)\\
&+c_i\left(1-\psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}-\texttt{AC-R}^0_{i,T}}{c_i}\right)-1/e\right)\Bigg)
\end{aligned}$$

Taking expectations over all sample paths completes the proof of Lemma 12. ☐

We now depart from the steps of Theorem 2 and derive a lower bound on the right hand side of (110) (and thus a lower bound on the sum of the pseudo-rewards) that no longer depends on $\texttt{AC-R}^{\text{EXT}}_{i,T}$ and $\texttt{AC-R}^0_{i,t}$.

Lemma 13 *For any instance $\mathcal{I}$, any sample path $\boldsymbol{\omega}$, and any opportunity i,*

$$\texttt{AC-R}^{\text{EXT}}_{i,T}+\texttt{AC-R}^0_{i,T}+\texttt{OPT}^{\text{INT}}_{i,T}\cdot\psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i-\texttt{AC-R}^{\text{EXT}}_{i,T}}\right)+c_i\left(1-\psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}-\texttt{AC-R}^0_{i,T}}{c_i}\right)-1/e\right)\geq(1-1/e)\texttt{OPT}_{i,T}, \tag{123}$$

where $\texttt{OPT}_{i,T}=\texttt{OPT}^{\text{EXT}}_{i,T}+\texttt{OPT}^{\text{INT}}_{i,T}$.

Proof of Lemma 13: We first note that the left hand side (LHS) of (123) is increasing in $\texttt{AC-R}^0_{i,T}$.

$$\begin{aligned}
\frac{\partial\text{ LHS}}{\partial\ \texttt{AC-R}^0_{i,T}}&=1+\psi'\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}-\texttt{AC-R}^0_{i,T}}{c_i}\right) &(124)\\
&=1-\exp\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}-\texttt{AC-R}^0_{i,T}}{c_i}-1\right) &(125)\\
&\geq 0 &(126)
\end{aligned}$$

The final inequality comes from noting that $\texttt{AC-R}^{\text{INT}}_{i,T} - \texttt{AC-R}^{0}_{i,T}$ cannot exceed the capacity c_i. Therefore, we can lower-bound the LHS by plugging in $\texttt{AC-R}^{0}_{i,T} = 0$ to yield

$$\text{LHS} \quad \geq \quad \texttt{AC-R}^{\text{EXT}}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T} \cdot \psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i - \texttt{AC-R}^{\text{EXT}}_{i,T}}\right) + c_i\left(1 - \psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i}\right) - 1/e\right) \tag{127}$$

There are now two cases to consider: (i) either $\texttt{AC-R}$ uses the same amount of external traffic as $\texttt{OPT}$ for opportunity i, or (ii) opportunity i reaches capacity under $\texttt{AC-R}$.[58]

In Case (i), we have

$$\begin{aligned}
\text{LHS} &\geq \texttt{OPT}^{\text{EXT}}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T} \cdot \psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i - \texttt{OPT}^{\text{EXT}}_{i,T}}\right) + c_i\left(1 - \psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i}\right) - 1/e\right) && (128)\\
&\geq \texttt{OPT}^{\text{EXT}}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T} \cdot \psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i - \texttt{OPT}^{\text{EXT}}_{i,T}}\right) + \texttt{OPT}_{i,T}\left(1 - \psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i}\right) - 1/e\right) && (129)\\
&= \texttt{OPT}^{\text{EXT}}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T} \cdot \psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i - \texttt{OPT}^{\text{EXT}}_{i,T}}\right) - (\texttt{OPT}^{\text{EXT}}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T}) \cdot \psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i}\right) + \texttt{OPT}_{i,T}\left(1 - 1/e\right) && (130)\\
&= (\texttt{OPT}^{\text{EXT}}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T}) \cdot \exp\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i} - 1\right) - \texttt{OPT}^{\text{INT}}_{i,T} \cdot \exp\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i - \texttt{OPT}^{\text{EXT}}_{i,T}} - 1\right) + \texttt{OPT}_{i,T}\left(1 - 1/e\right) && (131)\\
&= \exp\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i - \texttt{OPT}^{\text{EXT}}_{i,T}} - 1\right)\left((\texttt{OPT}^{\text{EXT}}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T}) \cdot \exp\left(\frac{-\texttt{AC-R}^{\text{INT}}_{i,T} \cdot \texttt{OPT}^{\text{EXT}}_{i,T}}{c_i(c_i - \texttt{OPT}^{\text{EXT}}_{i,T})}\right) - \texttt{OPT}^{\text{INT}}_{i,T}\right) + \texttt{OPT}_{i,T}\left(1 - 1/e\right) && (132)\\
&\geq \exp\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i - \texttt{OPT}^{\text{EXT}}_{i,T}} - 1\right)\left((\texttt{OPT}^{\text{EXT}}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T})\left(1 - \frac{\texttt{AC-R}^{\text{INT}}_{i,T} \cdot \texttt{OPT}^{\text{EXT}}_{i,T}}{c_i(c_i - \texttt{OPT}^{\text{EXT}}_{i,T})}\right) - \texttt{OPT}^{\text{INT}}_{i,T}\right) + \texttt{OPT}_{i,T}\left(1 - 1/e\right) && (133)\\
&= \exp\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i - \texttt{OPT}^{\text{EXT}}_{i,T}} - 1\right)\left(\texttt{OPT}^{\text{EXT}}_{i,T} - (\texttt{OPT}^{\text{EXT}}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T})\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T} \cdot \texttt{OPT}^{\text{EXT}}_{i,T}}{c_i(c_i - \texttt{OPT}^{\text{EXT}}_{i,T})}\right)\right) + \texttt{OPT}_{i,T}\left(1 - 1/e\right) && (134)\\
&\geq \exp\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i - \texttt{OPT}^{\text{EXT}}_{i,T}} - 1\right)\texttt{OPT}^{\text{EXT}}_{i,T}\left(1 - \frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i - \texttt{OPT}^{\text{EXT}}_{i,T}}\right) + \texttt{OPT}_{i,T}\left(1 - 1/e\right) && (135)\\
&\geq \texttt{OPT}_{i,T}\left(1 - 1/e\right) && (136)
\end{aligned}$$

Equality in (131) comes from applying the definition of the function ψ. Inequality (135) comes from noting that $\texttt{OPT}^{\text{EXT}}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T} \leq c_i$ and (136) comes from noting that in Case (i), where the amount of opportunity i's capacity filled by external traffic is the same under $\texttt{AC-R}$ and $\texttt{OPT}$, $\texttt{AC-R}^{\text{INT}}_{i,T} + \texttt{OPT}^{\text{EXT}}_{i,T} = \texttt{AC-R}^{\text{INT}}_{i,T} + \texttt{AC-R}^{\text{EXT}}_{i,T} \leq c_i$. This implies that $1 - \frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i - \texttt{OPT}^{\text{EXT}}_{i,T}} \geq 0$.

In Case (ii), where opportunity i reaches capacity under $\texttt{AC-R}$, we have

$$\begin{aligned}
\text{LHS} \quad &\geq \quad \texttt{AC-R}^{\text{EXT}}_{i,T} + \texttt{OPT}^{\text{INT}}_{i,T} \cdot \psi\left(\frac{c_i - \texttt{AC-R}^{\text{EXT}}_{i,T}}{c_i - \texttt{AC-R}^{\text{EXT}}_{i,T}}\right) + c_i\left(1 - \psi\left(\frac{c_i - \texttt{AC-R}^{\text{EXT}}_{i,T}}{c_i}\right) - 1/e\right) && (137)\\
&= \quad \texttt{AC-R}^{\text{EXT}}_{i,T} + c_i\left(1 - \psi\left(\frac{c_i - \texttt{AC-R}^{\text{EXT}}_{i,T}}{c_i}\right) - 1/e\right) && (138)\\
&= \quad \texttt{AC-R}^{\text{EXT}}_{i,T} - c_i \cdot \psi\left(1 - \frac{\texttt{AC-R}^{\text{EXT}}_{i,T}}{c_i}\right) + c_i\left(1 - 1/e\right) && (139)\\
&\geq \quad c_i\left(1 - 1/e\right) && (140)\\
&\geq \quad \texttt{OPT}_i\left(1 - 1/e\right) && (141)
\end{aligned}$$

[58] Based on Definition 1, $\texttt{OPT}$ will never use internal traffic to fill capacity that would otherwise be filled by external traffic. As a consequence, $\texttt{OPT}$ uses all external traffic for i (or fills opportunity i with external traffic) along each sample path. By our convention for external traffic, $\texttt{AC}$ will always *recommend* the volunteer's targeted opportunity i^*_t. However, if this opportunity has already reached capacity, the sign-up does not *fill* any capacity.

To establish (140), we note that the expression in (139) is non-decreasing in $\texttt{AC-R}^{\text{EXT}}_{i,T}$, as its derivative is given by $1-\exp(-\texttt{AC-R}^{\text{EXT}}_{i,T}/c_i)\geq 0$. Plugging in the smallest possible value for $\texttt{AC-R}^{\text{EXT}}_{i,T}$ (which is 0) yields (140).

This establishes (123) and completes the proof of Lemma 13. $\square$

By sequentially applying Lemmas 11, 12, and 13, we see that we can bound the expected amount of capacity filled under AC-R via the following inequalities:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\omega}}\left[\texttt{AC-R}\right] &\geq \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in[T]} L_t^{\mathcal{R}}+\sum_{i\in[n]} K_i^{\mathcal{R}}\right] && (142)\\
&\geq e^{-1/\underline{c}}\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} \texttt{AC-R}^{\text{EXT}}_{i,T}+\texttt{AC-R}^{0}_{i,T}+\texttt{OPT}^{\text{INT}}_{i,T}\cdot\psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}}{c_i-\texttt{AC-R}^{\text{EXT}}_{i,T}}\right)\right.\\
&\qquad\left.+c_i\left(1-\psi\left(\frac{\texttt{AC-R}^{\text{INT}}_{i,T}-\texttt{AC-R}^{0}_{i,T}}{c_i}\right)-1/e\right)\right] && (143)\\
&\geq e^{-1/\underline{c}}\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]}(1-1/e)\texttt{OPT}_i\right] && (144)\\
&= e^{-1/\underline{c}}(1-1/e)\mathbb{E}_{\boldsymbol{\omega}}\left[\texttt{OPT}\right] && (145)
\end{aligned}$$

This establishes a lower bound of $e^{-1/\underline{c}}(1-1/e)$ on the competitive ratio of AC-R, thereby completing the proof of Lemma 10. $\square$

Together with Lemma 9, this completes the proof of Proposition 4. $\square$

B.2. Proof of Proposition 5 (Section 5)

To prove Proposition 5, we use the approach of the proof of Theorem 2. In the following, we go through the main steps of that proof (as described in Section 4.4), and we provide detailed discussion of any adjustments needed to show that the result of Theorem 2 extends to this setting, which we henceforth refer to as the *cascade setting*. We emphasize that the cascade setting is a special case of the ranking setting where we can tailor our analysis to improve the bound (which, for the ranking setting, is given by Proposition 4).

To begin, we note that in the cascade setting, if the EFET is β, then the AC-R algorithm will fill at least a β fraction of capacity, as established in Lemma 9. We next prove the following additional lower bound on the competitive ratio of the AC-R algorithm in the cascade setting.

Lemma 14 *Let the smallest capacity be given by $\underline{c}$ and let the MCPR (given in Definition 4) be at most σ. Then, for any effective fraction of external traffic β, the competitive ratio of the* AC-R *algorithm in the cascade setting is at least z^* (as defined in (6)).*

This lemma is the analog (in the cascade setting) of Lemma 2, and to prove this result we follow the same three steps in the proof of Lemma 2, extended to this setting.

Step 1: Defining Pseudo-Rewards in the Cascade Setting

In the cascade setting, our notion of pseudo-rewards remains dependent on both the instance and the sample path. We extend our definition of a sample path such that $\boldsymbol{\omega}=\{\omega_1^v,\boldsymbol{\omega}_1^s\ldots,\omega_T^v,\boldsymbol{\omega}_T^s\}$ represents the realizations of random variables that govern both volunteer choices: the choice of which opportunity to *view* and the choice of which opportunity to *sign up for*, conditional on viewing. As volunteers' view decisions in this cascade setting are agnostic to the opportunity in each ranked position, we define ω_t^v as an integer between 1 and $K+1$, such that volunteer t views the opportunity that is ranked in position ω_t^v. We remind that the ranked subsets are of length at most K; hence, we use $\omega_t^v = K+1$ to indicate that the volunteer exits the platform (at any position) without viewing an opportunity. In general, a volunteer makes two random decisions for each considered position: whether to view and whether to exit if not viewing. However, $\omega_t^v \in [K+1]$ is sufficient information to fully specify the outcome of `AC-R` and `OPT`.

As in the base setting, we define $\boldsymbol{\omega}_t^s$ as a binary vector of length n, where the i^{th} component of $\boldsymbol{\omega}_t^s$ indicates whether volunteer t signs up for opportunity i, conditional on viewing opportunity i. We remark that, like all previous settings, given any fixed instance $\mathcal{I}$ and any fixed sample path $\boldsymbol{\omega}$, the output of `AC-R` and `OPT` are deterministic. We also remark that having $\omega_t^v \leq K$ is not a sufficient condition to ensure that the volunteer views an opportunity in that position, as it could be the case that the ranking presented to the volunteer was shorter than ω_t^v. In that case, again the volunteer does not view (or sign up for) any opportunity.

We further define the set $\mathcal{V}^{\mathcal{C}}$ as the set of internal traffic $t \in \mathcal{V}^{\text{INT}}$ for which t does not *view* an opportunity under `OPT`, along the given sample path $\boldsymbol{\omega}$. This expands on our definition of $\mathcal{V}^0$ in the base setting: as before, t is in $\mathcal{V}^{\mathcal{C}}$ if `OPT` does not recommend any opportunities. Now, we additionally have t in $\mathcal{V}^{\mathcal{C}}$ if the volunteer would view the opportunity ranked in position k (i.e., $\omega_t^v = k$) but `OPT` provides a ranking of length less than k. For instance, based on our assumption that the ranking provided is at most length K, volunteer t will be in $\mathcal{V}^{\mathcal{C}}$ if $\omega_t^v = K+1$. (We emphasize that the realization of ω_t^v is independent from the ranking provided by `OPT` for volunteer t.)

With this in mind, for the fixed instance $\mathcal{I}$ and the fixed sample path $\boldsymbol{\omega}$, we define the pseudo-rewards $L_t^{\mathcal{C}}$ for all $t \in [T]$ and $K_i^{\mathcal{C}}$ for all $i \in [n]$ according to the following:

$$L_t^{\mathcal{C}} = \begin{cases} \sum_{i\in[n]} \psi(\mathrm{FR}_{i,t-1}) \mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i], & t \in \mathcal{V}^{\text{EXT}} \cup \mathcal{V}^{\mathcal{C}} \\ \sum_{i\in[n]} \psi(\mathrm{FR}_{i,t-1}) \mathbb{1}[\xi_t(\vec{S}_t^{\texttt{OPT}}) = i], & t \in \mathcal{V}^{\text{INT}} \setminus \mathcal{V}^{\mathcal{C}} \end{cases} \tag{146}$$

$$K_i^{\mathcal{C}} = \sum_{t\in[T]} (1 - \psi(\mathrm{FR}_{i,t-1}))\, \mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i] \tag{147}$$

Step 2: Bounding the Value of `AC-R` in the Cascade Setting

This step of the proof involves two lemmas, which together establish a lower bound on the expected value of `AC-R` that depends (in part) on the expected value of `OPT`.

Lemma 15 *In the cascade setting, for any instance $\mathcal{I}$,*

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\texttt{AC-R}\right] \quad \geq \quad \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in[T]} L_t^{\mathcal{C}} + \sum_{i\in[n]} K_i^{\mathcal{C}}\right], \tag{148}$$

where $L_t^{\mathcal{C}}$ and $K_i^{\mathcal{C}}$ are defined in (146) and (147), respectively.

Proof of Lemma 15: This lemma is the analog (in the cascade setting) of Lemma 3 (proven in Appendix A.6.1). We follow the same algebraic steps, replicated below. As we later elaborate on, one particular inequality requires additional justification in the cascade setting.

$$\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC-R}] = \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^{\mathcal{C}}}\sum_{i\in[n]}\mathbb{1}\left[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})=i\right] + \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^{\mathcal{C}}}\sum_{i\in[n]}\mathbb{1}\left[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})=i|\right]\right] \tag{149}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]}\left(\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^{\mathcal{C}}}\psi(\text{FR}_{i,t-1})\mathbb{1}\left[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})=i\right] + \sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^{\mathcal{C}}}(1-\psi(\text{FR}_{i,t-1}))\mathbb{1}\left[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})=i\right]\right.\right.$$
$$\left.\left. + \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^{\mathcal{C}}}\psi(\text{FR}_{i,t-1})\mathbb{1}\left[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})=i\right] + \sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^{\mathcal{C}}}(1-\psi(\text{FR}_{i,t-1}))\mathbb{1}\left[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})=i\right]\right)\right] \tag{150}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^{\mathcal{C}}}\psi(\text{FR}_{i,t-1})\mathbb{1}\left[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})=i\right]\right] + \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^{\mathcal{C}}}L_t^{\mathcal{C}} + \sum_{i\in[n]}K_i^{\mathcal{C}}\right] \tag{151}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^{\mathcal{C}}}\psi(\text{FR}_{i,t-1})\mathbb{1}\left[\xi_t(\vec{S}_t^{\texttt{AC-R}})=i\right]\right] + \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^{\mathcal{C}}}L_t^{\mathcal{C}} + \sum_{i\in[n]}K_i^{\mathcal{C}}\right] \tag{152}$$

$$\geq \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^{\mathcal{C}}}\psi(\text{FR}_{i,t-1})\mathbb{1}\left[\xi_t(\vec{S}_t^{\texttt{OPT}})=i\right]\right] + \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in\mathcal{V}^{\text{EXT}}\cup\mathcal{V}^{\mathcal{C}}}L_t^{\mathcal{C}} + \sum_{i\in[n]}K_i^{\mathcal{C}}\right] \tag{153}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in[T]}L_t^{\mathcal{C}} + \sum_{i\in[n]}K_i^{\mathcal{C}}\right] \tag{154}$$

All steps are algebraic except for (152) and Line (153). To establish the former, we will show that $\sum_{i\in[n]}\psi(\text{FR}_{i,t-1})\mathbb{1}[\xi_t(\vec{S}_t^{\texttt{AC-R}}) = i] = \sum_{i\in[n]}\psi(\text{FR}_{i,t-1})\mathbb{1}[\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = i]$. We consider two cases. First, if $\text{FR}_{\xi_t(\vec{S}_t^{\texttt{AC-R}}),t-1} < 1$, then $\xi_t(\vec{S}_t^{\texttt{AC-R}}) = \tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}})$ and the equality holds. Alternatively, if $\text{FR}_{\xi_t(\vec{S}_t^{\texttt{AC-R}}),t-1} = 1$, then $\tilde{\xi}_t(\vec{S}_t^{\texttt{AC-R}}) = 0$ and $\psi(\text{FR}_{\xi_t(\vec{S}_t^{\texttt{AC-R}}),t-1}) = 0$. Thus, both summations equal 0, and the equality holds.

Establishing (153) in the cascade setting requires more care, as the set $\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^{\mathcal{C}}$ only includes volunteers that *viewed* an opportunity under OPT, and whether or not a volunteer views an opportunity under OPT depends on the ranking provided by OPT. To that end, it is sufficient to show the following inequality holds for all $t\in\mathcal{V}^{\text{INT}}$, where we define $\boldsymbol{\omega}_{-t}$ as a sample path excluding the realizations governing the decisions of volunteer t (i.e., ω_t^v and $\boldsymbol{\omega}_t^s$).

$$\mathbb{E}_{\omega_t^v,\boldsymbol{\omega}_t^s}\left[\sum_{i\in[n]}\psi(\text{FR}_{i,t-1})\mathbb{1}\left[\xi_t(\vec{S}_t^{\texttt{AC-R}})=i\right]\mathbb{1}\left[t\notin\mathcal{V}^{\mathcal{C}}\right] \mid \boldsymbol{\omega}_{-t}\right]$$
$$\geq \mathbb{E}_{\omega_t^v,\boldsymbol{\omega}_t^s}\left[\sum_{i\in[n]}\psi(\text{FR}_{i,t-1})\mathbb{1}\left[\xi_t(\vec{S}_t^{\texttt{OPT}})=i\right]\mathbb{1}\left[t\notin\mathcal{V}^{\mathcal{C}}\right] \mid \boldsymbol{\omega}_{-t}\right] \tag{155}$$

Applying the tower property of expectations would then establish the validity of (153).

To show that (155) holds, we first take advantage of the fact that, in the cascade setting, the probability of viewing an opportunity under any algorithm (including OPT) depends only on the *length* of the ranking provided by that algorithm, and not on the identity and ordering of the opportunities in the ranking.

To be precise, we make the following claim:

Claim 7 *In the cascade setting, for a fixed instance $\mathcal{I}$ and a fixed sample path $\boldsymbol{\omega}$, for any volunteer $t\in\mathcal{V}^{\text{INT}}$ there is a position $k_t^*(\boldsymbol{\omega}_{-t})$ such that $t\in\mathcal{V}^{\text{INT}}\setminus\mathcal{V}^{\mathcal{C}}$ if and only if $\omega_t^v \leq k_t^*(\boldsymbol{\omega}_{-t})$.*

Proof of Claim 7: Recall our convention that OPT recommends the optimal ranking which is of shortest length (breaking ties in favor of the lexicographically smallest such subset with respect to its indices). This convention – in combination with the fact that the probability of a volunteer viewing an opportunity is decreasing in the opportunity's rank in the cascade setting – ensures an important property of OPT: *if* OPT *ranks a (non-dummy) opportunity in position k, it will also rank (non-dummy) opportunities in positions* 1 *through* $k-1$. To see why, suppose that this is not the case. Moving the last-ranked opportunity up to an open position (i.e., a position occupied by a dummy opportunity) shortens the ranking, and doing so weakly increases the amount of filled capacity under OPT. Thus, OPT should have recommended this ranking, which establishes a contradiction.

As a consequence, let k_t^* denote the length of the ranking provided to volunteer $t \in \mathcal{V}^{\text{INT}}$ by OPT along sample path $\boldsymbol{\omega}$. If t views an opportunity under OPT, then $\omega_t^v \leq k_t^*$. The converse also holds.

We remark that the length of this ranking (i.e., k_t^*) depends only on the number of opportunities with remaining capacity for internal traffic at time $t-1$ (for which volunteer t has positive conversion probability). This set of opportunities is not a function of the realizations ω_t^v and $\boldsymbol{\omega}_t^s$. □

In light of Claim 7, we can rewrite (155) as follows, using $\mathbb{P}^{\mathcal{C}}$ to denote the probability distribution associated with the realizations ω_t^v, which depends only on the parameters of the opportunity-agnostic cascade model. Furthermore, we use $\vec{S}_t^{\texttt{AC-R}}(k)$ (resp., $\vec{S}_t^{\texttt{OPT}}(k)$) to denote the opportunity ranked in position k under AC-R (resp. OPT).

$$
\begin{aligned}
&\mathbb{E}_{\omega_t^v,\boldsymbol{\omega}_t^s}\left[\sum_{i\in[n]}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}\left[\xi_t(\vec{S}_t^{\texttt{AC-R}})=i\right]\mathbb{1}\left[\omega_t^v\leq k_t^*(\boldsymbol{\omega}_{-t})\right]\;\middle|\;\boldsymbol{\omega}_{-t}\right]\\
&\qquad=\sum_{k\in[k_t^*(\boldsymbol{\omega}_{-t})]}\mathbb{P}^{\mathcal{C}}\left[\omega_t^v=k\;\middle|\;\omega_t^v\leq k_t^*(\boldsymbol{\omega}_{-t})\right]\mathbb{P}^{\mathcal{C}}\left[\omega_t^v\leq k_t^*(\boldsymbol{\omega}_{-t})\right]\mu_{\vec{S}_t^{\texttt{AC-R}}(k),t}\psi(\mathrm{FR}_{\vec{S}_t^{\texttt{AC-R}}(k),t-1}) &(156)\\
&\qquad=\sum_{k\in[k_t^*(\boldsymbol{\omega}_{-t})]}\mathbb{P}^{\mathcal{C}}\left[\omega_t^v=k\right]\mu_{\vec{S}_t^{\texttt{AC-R}}(k),t}\psi(\mathrm{FR}_{\vec{S}_t^{\texttt{AC-R}}(k),t-1}) &(157)\\
&\qquad\geq\sum_{k\in[k_t^*(\boldsymbol{\omega}_{-t})]}\mathbb{P}^{\mathcal{C}}\left[\omega_t^v=k\right]\mu_{\vec{S}_t^{\texttt{OPT}}(k),t}\psi(\mathrm{FR}_{\vec{S}_t^{\texttt{OPT}}(k),t-1}) &(158)\\
&\qquad=\sum_{k\in[k_t^*(\boldsymbol{\omega}_{-t})]}\mathbb{P}^{\mathcal{C}}\left[\omega_t^v=k\;\middle|\;\omega_t^v\leq k_t^*(\boldsymbol{\omega}_{-t})\right]\mathbb{P}^{\mathcal{C}}\left[\omega_t^v\leq k_t^*(\boldsymbol{\omega}_{-t})\right]\mu_{\vec{S}_t^{\texttt{OPT}}(k),t}\psi(\mathrm{FR}_{\vec{S}_t^{\texttt{OPT}}(k),t-1}) &(159)\\
&\qquad=\mathbb{E}_{\omega_t^v,\boldsymbol{\omega}_t^s}\left[\sum_{i\in[n]}\psi(\mathrm{FR}_{i,t-1})\mathbb{1}\left[\xi_t(\vec{S}_t^{\texttt{OPT}})=i\right]\mathbb{1}\left[\omega_t^v\leq k_t^*(\boldsymbol{\omega}_{-t})\right]\;\middle|\;\boldsymbol{\omega}_{-t}\right] &(160)
\end{aligned}
$$

We note that equality in (157) and (159) follow from the rules of conditional probability, as for any $k \leq k_t^*(\boldsymbol{\omega}_{-t})$, we have $\mathbb{P}^{\mathcal{C}}\left[\omega_t^v=k \mid \omega_t^v\leq k_t^*(\boldsymbol{\omega}_{-t})\right]=\frac{\mathbb{P}^{\mathcal{C}}[\omega_t^v=k]}{\mathbb{P}^{\mathcal{C}}[\omega_t^v\leq k_t^*(\boldsymbol{\omega}_{-t})]}$. All that remains is to prove that (158) holds, which we do via the following claim:

Claim 8 *Let $\vec{S}_t^{\texttt{AC-R}}$ be the ranking presented by* AC-R *to volunteer $t \in \mathcal{V}^{\text{INT}}$, as given by* (12)*. Then, in the cascade setting, $\vec{S}_t^{\texttt{AC-R}}$ also satisfies the following condition for any $k' \leq K$:*

$$\vec{S}_t^{\texttt{AC-R}} \in \operatorname{argmax}_{\vec{S}} \sum_{k\in[k']}\mathbb{P}^{\mathcal{C}}\left[\omega_t^v=k\right]\mu_{\vec{S}(k),t}\psi(\mathrm{FR}_{\vec{S}(k),t-1}).$$

Proof of Claim 8 Applying the optimality condition of the AC-R algorithm (see (12)) to the cascade setting, we see that

$$\vec{S}_t^{\texttt{AC-R}} \in \operatorname{argmax}_{\vec{S}} \sum_{k\in[K]} \mathbb{P}^{\mathcal{C}}\left[\omega_t^v = k\right]\mu_{\vec{S}(k),t}\psi(\mathrm{FR}_{\vec{S}(k),t-1}).$$

To prove Claim 8, we need to show that that AC-R continues to satisfy this optimality condition when considering the sum over the first k' terms, for any $k' \le K$. In the cascade setting, the view probability depends only on an opportunity's position, and $\mathbb{P}[\omega_t^v = k] = \nu_t\big((1-\nu_t)(1-q_t)\big)^{k-1}$ is decreasing in the position k. Therefore, the AC-R algorithm will rank opportunities in descending order of $\mu_{i,t}\psi(\mathrm{FR}_{i,t-1})$ (breaking ties in favor of the lowest-indexed opportunity), until it exhausts the maximum list size K. To see why, suppose $\mu_{i,t}\psi(\mathrm{FR}_{i,t-1}) > \mu_{j,t}\psi(\mathrm{FR}_{j,t-1})$, but opportunity i is ranked after opportunity j in $\vec{S}_t^{\texttt{AC-R}}$. In that case, switching opportunity i and opportunity j in $\vec{S}_t^{\texttt{AC-R}}$ would strictly increase the objective that AC-R is optimizing for, which represents a contradiction.

By an identical argument, ranking opportunities in descending order of $\mu_{i,t}\psi(\mathrm{FR}_{i,t-1})$ also maximizes

$$\sum_{k\in[k']} \mathbb{P}^{\mathcal{C}}\left[\omega_t^v = k\right]\mu_{\vec{S}(k),t}\psi(\mathrm{FR}_{\vec{S}(k),t-1}).$$

Therefore, AC-R also satisfies this optimality condition for any k', which completes the proof of Claim 8. □

Together, Claims 7 and 8 prove that (153) holds. This completes the proof of Lemma 15. □

Lemma 16 *In the cascade setting, for any instance $\mathcal{I}$,*

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{t\in[T]} L_t^{\mathcal{C}} + \sum_{i\in[n]} K_i^{\mathcal{C}}\right] \quad\ge\quad e^{-1/c}\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} \texttt{AC-R}_{i,T}^{\mathrm{EXT}} + \texttt{AC-R}_{i,T}^{\mathcal{C}} + \texttt{OPT}_{i,T}^{\mathrm{INT}}\cdot\psi\left(\frac{\texttt{AC-R}_{i,T}^{\mathrm{INT}}}{c_i - \texttt{AC-R}_{i,T}^{\mathrm{EXT}}}\right)\right.$$
$$\left. + c_i\left(1-\psi\left(\frac{\texttt{AC-R}_{i,T}^{\mathrm{INT}} - \texttt{AC-R}_{i,T}^{\mathcal{C}}}{c_i}\right) - 1/e\right)\right], \tag{161}$$

where $L_t^{\mathcal{C}}$ and $K_i^{\mathcal{C}}$ are defined in (146) and (147), respectively.

Lemma 16 is the analog (in the cascade setting) of Lemma 12. The proof of this lemma immediately follows by taking identical steps as in the proof of Lemma 12. (As the proof is algebraic and holds along each sample path, the distinction between $\mathcal{V}^0$ and $\mathcal{V}^{\mathcal{C}}$ does not impact the result.) We omit these details for the sake of brevity.

Step 3: Bounding the Competitive Ratio of AC-R in the Cascade Setting

The final step of the proof involves the use of the instance-specific mathematical program (MP) (see Table 1), which helps establish a lower bound on the competitive ratio of AC-R in the cascade setting.

Lemma 17 *In the cascade setting, for any instance $\mathcal{I}$, the ratio between the expected value of* AC-R *(i.e., $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC-R}]$) and the expected value of* OPT *(i.e., $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{OPT}]$) on instance $\mathcal{I}$ is at least the value of* (MP).

Proof of Lemma 17: Lemma 17 is the analog (in the cascade setting) of Lemma 5, and our proof follows a similar approach. To prove Lemma 17, we consider the following candidate solution:

$$x_{1,i,\boldsymbol{\omega}} = \texttt{AC-R}^{\text{EXT}}_{i,T}, \qquad x_{2,i,\boldsymbol{\omega}} = \texttt{AC-R}^{\text{INT}}_{i,T}, \qquad x_{3,i,\boldsymbol{\omega}} = \texttt{AC-R}^{\mathcal{C}}_{i,T},$$
$$y_{1,i,\boldsymbol{\omega}} = \texttt{OPT}^{\text{EXT}}_{i,T}, \qquad y_{2,i,\boldsymbol{\omega}} = \texttt{OPT}^{\text{INT}}_{i,T}, \qquad z = \frac{\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC-R}]}{\mathbb{E}_{\boldsymbol{\omega}}[\texttt{OPT}]}$$

Such a solution has an objective value equal to the ratio $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC-R}]/\mathbb{E}_{\boldsymbol{\omega}}[\texttt{OPT}]$ in (MP), and by construction it satisfies all constraints. The first five constraints hold by exactly the same rationale described in the proof of Lemma 5 (see Appendix A.6.3).

To see that the sixth constraint is satisfied, let us fix a sample path $\boldsymbol{\omega}$ and an opportunity i. The total amount of opportunity i's capacity filled by `AC-R` in periods $t \in \mathcal{V}^{\text{INT}} \setminus \mathcal{V}^{\mathcal{C}}$ is given by $x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}}$, while the total amount of opportunity i's capacity filled by `OPT` in periods $t \in \mathcal{V}^{\text{INT}} \setminus \mathcal{V}^{\mathcal{C}}$ is given by $y_{2,i,\boldsymbol{\omega}}$. Furthermore, for all $t \in \mathcal{V}^{\text{INT}} \setminus \mathcal{V}^{\mathcal{C}}$, volunteer t views an opportunity under `OPT`, which means it fills a unit of capacity with probability at least $\underline{\mu}_t$. For the same volunteer t, `AC-R` will fill a unit of capacity with probability at most $\bar{\mu}_t$. As a consequence, $x_{2,i,\boldsymbol{\omega}} - x_{3,i,\boldsymbol{\omega}} \leq \sigma y_{2,i,\boldsymbol{\omega}}$, or equivalently, $x_{2,i,\boldsymbol{\omega}} \leq \sigma y_{2,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}}$.

Based on the constructed values of $\vec{x}, \vec{y}$, and z, as well as the upper bound on $x_{2,i,\boldsymbol{\omega}}$ identified above,

$$\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{1,i,\boldsymbol{\omega}}\right] = z \cdot \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\right] - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{2,i,\boldsymbol{\omega}}\right] \tag{162}$$
$$\geq z \cdot \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\right] - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} \sigma \cdot y_{2,i,\boldsymbol{\omega}} + x_{3,i,\boldsymbol{\omega}}\right] \tag{163}$$
$$= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}}\right] - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} (1-z)\cdot y_{1,i,\boldsymbol{\omega}} + (\sigma - z)\cdot y_{2,i,\boldsymbol{\omega}}\right] - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{3,i,\boldsymbol{\omega}}\right] \tag{164}$$
$$\geq \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}}\right] - (\sigma - z)\cdot \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}} + y_{2,i,\boldsymbol{\omega}}\right] - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{3,i,\boldsymbol{\omega}}\right] \tag{165}$$
$$\geq \beta \sum_{i\in[n]} c_i - (\sigma - z)\sum_{i\in[n]} c_i - \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} x_{3,i,\boldsymbol{\omega}}\right]. \tag{166}$$

Inequality (165) uses the fact that $\sigma \geq 1$. The final inequality uses the fact that $\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{i\in[n]} y_{1,i,\boldsymbol{\omega}}\right] = \beta \sum_{i\in[n]} c_i$ based on the definitions of the optimal clairvoyant algorithm `OPT` and the EFET β (see Definitions 1 and 2). This final inequality establishes that our candidate solution respects constraint (vi).

The fact that the candidate solution satisfies the seventh (and final constraint) follows by applying Lemmas 15 and 16 from Step 2. This completes the proof of Lemma 17. □

As established in Lemma 6 (proven in Appendix A.6.4), the optimal value of (MP) is at least z^* (as defined in (6)). Therefore, we have shown a lower bound on the ratio $\mathbb{E}_{\boldsymbol{\omega}}[\texttt{AC-R}]/\mathbb{E}_{\boldsymbol{\omega}}[\texttt{OPT}]$ in the cascade setting for any instance $\mathcal{I} \in \mathcal{I}_\beta$, where the bound depends on only the EFET β, the minimum capacity $\underline{c}$, and the MCPR σ.

Taken as a whole, these three steps prove Lemma 14, namely, that z^* is a lower bound on the competitive ratio of the `AC-R` algorithm in the cascade setting. Thus, in combination with our observation that the competitive ratio is lower-bounded by β, we have shown that the competitive ratio of the `AC-R` algorithm is at least $f(\beta, \underline{c}, \sigma)$, as defined in the statement of Theorem 2. □

Appendix C: Additional Details on Case Study

C.1. Data Availability

Through our partnership with VolunteerMatch, we have access to the following sources of detailed data on their platform:

1. Volunteer Match's back-end database that provides opportunity-level data on characteristics such as their posting dates, locations (in-person or virtual), timings (specific dates/times or a flexible schedule), capacities (i.e., the number of volunteers needed), and the cause(s) the organization supports (out of a list of thirty including LGBTQ, seniors, hunger, etc.). To ensure consistent data quality and accuracy, we limit our analysis to virtual opportunities active between August 2020- March 2021 for which we have precise data on capacity (i.e., those that request a number of volunteers between 1 and 20). We focus on virtual opportunities, as these opportunities do not have compatibility that depends on the proximity of a volunteer to an opportunity.

2. Google Analytics (GA) data that details user behavior on the site. GA provides session-level information for all devices accessing the website, allowing us to understand the different ways users access the site, whether this source of access affects search behavior, and whether different kinds of opportunities are more likely to be viewed or more likely to be signed-up for, conditional on being viewed. We have access to data for activity between August 2020- March 2021 for devices from New York City, Miami, Austin, Alaska, Maine, Montana, Vermont, and West Virginia. Opportunities appearing in our GA dataset are those that were viewed at least once by one or more of these devices.

For the window between August 2020-March 2021, we combine these datasets, which allows us to directly determine the arrival order of internal and external traffic, and to estimate opportunity-level conversion probabilities conditional on a view.

C.2. Set of Opportunities and Robustness

In Section 6, in order to calculate $\overline{\texttt{OPT}}$, we focus on a simple random sample of 100 opportunities from the 10,737 virtual opportunities that appear in our GA dataset between August 2020 and March 2021 for which we have precise data on capacity. In this section, we broaden our focus to the full set of 10,737 opportunities. First, we note that the EFET of our sample is 19% compared to the EFET of 17% for the full dataset. Though it is intractable to calculate $\overline{\texttt{OPT}}$ for the full dataset, we simulate AC and compare it to another algorithm (SCP, defined in Section 6.2) over different instance sizes. In Figure 7a we show the value of AC increases linearly with instance size, while the value of SCP increases at a slightly lower rate, resulting in even better relative performance for AC for larger instance sizes.[59] This highlights the importance of accounting for conversion probabilities: the value of SCP scales sublinearly because its recommendations predominantly consist of opportunities with frequent action but low conversion probability.

[59] The values of AC and SCP are averaged over sufficient numbers of samples, dependent on the instance size.

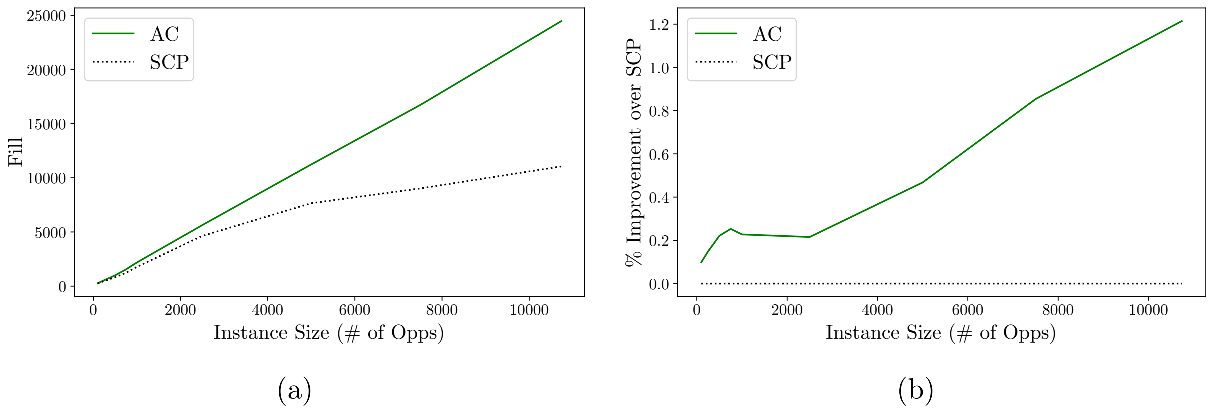


Figure 7 **(a) The value of AC and SCP over different instance sizes and (b) the percent improvement of AC over SCP over different instance sizes.**

C.3. Arrival Sequence and Volunteer Conversion Probabilities

In this section, we provide additional details on how we construct the instance used in Section 6. Based on volunteer activity, we estimate that there are $T = 11,345$ website visitors for our subset of 100 opportunities. To arrive at this estimate, we first observe a subset of views for each opportunity, which we then scale upward proportionally (by a factor of 5) to account for additional unobserved nation-wide traffic that views each opportunity. The scaling factor of 5 was determined using VM's back-end data, as the Google Analytics data only contains data from volunteers in a subset of major cities and rural regions, listed in Appendix C.1, meaning we only observe session-level data for approximately 20% of website traffic. We further augment the number of arriving volunteers to account for the 45% of internal visitors who do not view any opportunity.

To generate the arrival sequence, we preserve the order of the traffic pattern observed in the session-level data after appropriate scaling. For example, if volunteer 1 is external traffic who signs up for opportunity a and volunteer 2 is internal traffic in the dataset, then in our simulation, the first five volunteers will all be external traffic that will sign-up for opportunity a and the next five will be internal traffic for which we generate pair-specific conversion probabilities. We note that volunteers six through ten may not have the same conversion probabilities, but rather are each drawn independently from the distribution described in Section 6. We do this because while a sign-up from external traffic could only have occurred in one way, there are many counterfactual options for internal traffic that we cannot observe.

For each compatible opportunity, we estimate $\mu_{i,t}$ via a comprehensive logistic regression, using observable opportunity characteristics as explanatory variables, including the opportunity's cause and the conversion probabilities of views on the first day that an opportunity is viewed within our time-frame, as we have access to both views and sign-ups from GA. Though our model does not account for learning, we include the latter in order to generate the most accurate estimates of conversion probabilities; we find that there is limited value to learning past the first day. The result is presented in Table 2

Opportunity Characteristics	Internal Conversion Probabilities Odds-Ratio	Std. Error
intercept	-2.96***	0.17
first_views_outcome	2.85***	0.08
women	-0.25	0.16
justice	-0.19	0.25
politics	0.06	0.34
homeless	-0.03	0.20
health	-0.09	0.13
seniors	0.16	0.17
environment	-0.27	0.16
crisis	-0.11	0.17
education	-0.09	0.11
disabilities	-0.06	0.17
sports	0.41	0.26
children	-0.09	0.11
advocacy	0.08	0.11
community	-0.08	0.09
employment	-0.08	0.21
emergency	0.05	0.31
international	0.00	0.15
animals	0.00	0.18
lgbtq	0.15	0.30
race	0.24	0.23
arts	-0.03	0.14
faith_based	-0.20	0.26
computers	-0.10	0.13
board	0.08	0.19
hunger	0.10	0.21
media	-0.19	0.16
immigrants	-0.19	0.20
disaster	-0.24	0.28
veterans	0.03	0.22
Log- Likelihood	-1,385.1	

Table 2 Logit Regression model to predict internal and external traffic conversion probabilities. The regression uses 10,737 observations. P-values: < 0.01 *, < 0.05 **, < 0.10 *.**

C.4. Formal Definition of $\overline{\texttt{OPT}}$

Here we present our definition of $\overline{\texttt{OPT}}$, which is the solution to the following program, where $f(k)$ represents the probability of viewing a particular rank $k \in [K]$. Specifically, given the parameters of the opportunity-agnostic cascade model given in Section 6, $f(k) = 0.3 + ((1-0.3)\cdot(1-0.24))^k$.

$$
\begin{aligned}
\max_{\vec{x}} \quad & \sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{INT}}}\sum_{k\in[K]} \mu_{i,t}x_{i,t,k}f(k) + \sum_{i\in[n]}\sum_{t\in\mathcal{V}^{\text{EXT}}} \mu_{i^*_t,t} \\
\text{subject to} \quad & \sum_{t\in\mathcal{V}^{\text{INT}}}\sum_{k\in[K]} \mu_{i,t}x_{i,t,k}f(k) + \sum_{t\in\mathcal{V}^{\text{EXT}}} \mu_{i^*_t,t} \le c_i \qquad \forall i\in[n] && (1)\\
& \sum_{i\in[n]} x_{i,t,k} \le 1 \qquad \forall k\in[K], t\in\mathcal{V}^{\text{INT}} && (2)\\
& \sum_{k\in[K]} x_{i,t,k} \le 1 \qquad \forall i\in[n], t\in\mathcal{V}^{\text{INT}} && (3)\\
& 0\le x_{i,t,k}\le 1; \qquad \forall i\in[n], t\in\mathcal{V}^{\text{INT}}, k\in[K] && (4)
\end{aligned}
$$

This program uses the set of variables $\vec{x}\in\mathbb{R}^{n\times|\mathcal{V}^{\text{INT}}|\times K}$. For the case study of Section 6, we choose $K = 3$ for tractable computation. This linear program is a deterministic fractional matching. As formalized in the proposition below, the optimal value of $\overline{\texttt{OPT}}$ is thus an upper bound on the expected value of $\texttt{OPT}$.

Proposition 7 $\overline{\texttt{OPT}}$ *is an upper bound on the expected value of* $\texttt{OPT}$.

Proof of Proposition 7: To prove this, we will show that there exists a feasible solution in $\overline{\texttt{OPT}}$ that achieves the expected value of $\texttt{OPT}$. Let $\hat{x}$ be the ex-ante probability that opportunity i is assigned to position k in the ranking shown to volunteer $t\in\mathcal{V}^{\text{INT}}$ under $\texttt{OPT}$. To see that $\hat{x}$ is a feasible solution in $\overline{\texttt{OPT}}$, note that (i) $\texttt{OPT}$ can only fill each opportunity to capacity, (ii) at most one opportunity can appear in each position $k\in[K]$, (iii) each opportunity $i\in[n]$ can appear in a ranking at most once, and (iv) a valid probability must be between 0 and 1. By construction, the value of $\overline{\texttt{OPT}}$ under $\hat{x}$ is equal to the expected value of $\texttt{OPT}$. Since any feasible value of the linear program must be less than or equal to the optimal value of the linear program, we see that the expected value of $\texttt{OPT}$ must be less than or equal to the solution of $\overline{\texttt{OPT}}$. $\square$