Proportionally Fair Online Allocation of Public Goods with Predictions

Siddhartha Banerjee\(^1\), Vasilis Gkatzelis\(^2\), Safwan Hossain\(^3\), Billy Jin\(^1\), Evi Micha\(^4\) and Nisarg Shah\(^4\)

\(^1\)Cornell University\n\(^2\)Drexel University\n\(^3\)Harvard University\n\(^4\)University of Toronto

sbanerjee@cornell.edu, gkatz@drexel.edu, shossain@fas.harvard.edu, bzj3@cornell.edu, emicha@cs.toronto.edu, nisarg@cs.toronto.edu

Abstract

We design online algorithms for the fair allocation of public goods to a set of \(N\) agents over a sequence of \(T\) rounds and focus on improving their performance using predictions. In the basic model, a public good arrives in each round, and every agent reveals their value for it upon arrival. The algorithm must irrevocably decide the investment in this good without exceeding a total budget of \(B\) across all rounds. The algorithm can utilize (potentially noisy) predictions of each agent's total value for all public goods. The algorithm's performance is measured using a proportional fairness objective, which formally demands that every group of agents be rewarded proportional to its size and the cohesiveness of its preferences. We show that no algorithm can achieve better than \(\Theta(T/B)\) proportional fairness without predictions. With reasonably accurate predictions, the situation improves significantly, and \(\Theta(\log(T/B))\) proportional fairness is achieved. We also extend our results to a general setting wherein a batch of \(L\) public goods arrive in each round and \(O(\log(\min(N,L) \cdot T/B))\) proportional fairness is achieved. Our exact bounds are parameterized as a function of the prediction error, with performance degrading gracefully with increasing errors.

1 Introduction

In classic online algorithms, the input is presented in stages and the algorithm needs to make irrevocable decisions at each stage without knowing the input from future stages. Its performance, called the competitive ratio, is measured by comparing the worst-case ratio of the achieved solution quality to the optimal solution quality in hindsight [Borodin and El-Yaniv, 2005]. The uncertainty regarding the future often forces such algorithms to make overly cautious decisions, resulting in unappealing competitive ratios.

An emerging line of research asks whether one can utilize the progress in machine learning (ML) to augment online algorithms with machine-learned predictions regarding the future [Mitzenmacher and Vassilvitskii, 2021]. Ideally, one would hope for bicriteria guarantees, ensuring that the competitive ratio is improved when the predictions are reasonably accurate but remaining robust even when they are not. More generally, one can hope to express the competitive ratio of the online algorithm in terms of the error in the predictions. This powerful paradigm has received significant attention, including from the machine learning community, for problems such as caching [Lykouris and Vassilvitskii, 2018; Rohatgi, 2020; Jiang et al., 2020], the secretary problem [Dütting et al., 2021; Antoniadis et al., 2020a; Antoniadis et al., 2020b], scheduling [Lattanzi et al., 2020], the ski rental problem [Purohit et al., 2018; Wang et al., 2020], set cover [Bamas et al., 2020], and other problems [Almanza et al., 2021; Antoniadis et al., 2021; Dinitz et al., 2021].

However, these works are all limited to single-agent decision-making problems. Recently, Banerjee et al. [2022a] applied this paradigm to design online algorithms for a multi-agent resource allocation problem, in which a set of private goods (which can only be allocated to and enjoyed by a single agent) need to be divided amongst a group of agents in a fair manner. Using the Nash welfare from bargaining theory [Nash, 1950] as their notion of fairness, Banerjee et al. [2022a] show that predictions about agents' total value can be utilized to achieve significantly improved approximations.

The solutions proposed by Banerjee et al. [2022a], however, do not capture resource allocation settings involving public goods, i.e., goods whose benefit can be enjoyed by multiple agents (e.g., a highway or a park). In many important problems, like participatory budgeting, committee selection, or shared memory allocation, some scarce resources need to be dedicated to make each public good available, and an algorithm needs to decide which goods to invest in, aiming to make the agents happy. In participatory budgeting, it is common to include projects such as installation of bike racks or public restrooms, road maintenance, or renovation of public school facilities, where partial investment is also possible and can lead to meaningful progress towards the ambitious goals.

Fairness in these settings is often captured by notions like the Nash welfare and the core [Foley, 1970]. Yet, despite the significance of its applications, only a few papers have successfully studied the fair allocation of public goods [Fain et al., 2016; Kunjir et al., 2017; Friedman et al., 2019; Conitzer et al., 2017; Fain et al., 2018; Peters et al., 2020;
Munagala et al., 2021} (relative to the extensive literature on private goods; see [Brandt et al., 2016]), even fewer have provided online algorithms for this problem [Freeman et al., 2017], and none utilize predictions. We address this gap by designing online algorithms for fair allocation of public goods using predictions about how agents value the goods. The key research questions we address are:

**How can we allocate public goods online in a fair manner? Can predictions about agent preferences help improve fairness guarantees?**

### 1.1 Our Results

We study online algorithms for the fair allocation of public goods arriving over a sequence of $T$ rounds based on the preferences of $N$ agents. In the basic model a new public good arrives in each round $t$ (referred to as good $t$), and the algorithm learns the value $v_{i,t}$ of each agent $i$ for this good. Using this information, the algorithm needs to make an irrevocable decision regarding an amount $x_t \in [0, 1]$ to invest in this good. Each agent $i$ then receives value $v_{i,t} \cdot x_t$ from this investment, with the total utility of an agent being the sum of the values gained across rounds. While the algorithm would like to increase the $x_t$’s as much as possible, it is limited by a total budget constraint: $\sum_t x_t \leq B$, where $B$ is given. In formally, the budget constraint encourages the algorithm to invest more in goods that are highly valued by many agents. However, this is challenging since the agents’ values for future goods are unknown. To deal with this uncertainty, we assume the algorithm has access to predictions regarding the total value of each agent $i$: this prediction $\hat{V}_i$ provides an estimate on the total value $V_i = \sum_t v_{i,t}$ of agent $i$ for all goods to arrive. We discuss a bit more on why predictions can be reasonable to obtain in practice in Section 4.1.

We focus on a quantitative fairness objective, called proportional fairness (Definition 1), which is stronger than previously considered objectives of the Nash welfare and the core (see Section 2.1). Lower objective values indicate better fairness, with 1 indicating perfect proportional fairness.

In Section 3, as a warm-up, we consider the setting where agent values are binary (i.e., $v_{i,t} \in \{0, 1\}$ for all agents $i$ and goods $t$) and the budget is $B = 1$. Binary values correspond to approval voting – where agents either like a good or not; the unit budget forces $\sum_t x_t \leq 1$, meaning that $x_t$ can also be viewed as the fraction of an available resource invested in good $t$. For this special case, we show that it is already possible to achieve $O(\log N)$ proportional fairness without using any predictions and this is optimal even if the algorithm had access to perfect predictions ($\hat{V}_i = V_i$ for each agent $i$).

In Section 4, we consider general values and budget, and show that without predictions, no algorithm can achieve $o(T/B)$ proportional fairness even when there is $N = 1$ agent. In contrast, by using the predictions, we can achieve an exponential improvement to $O(\log(T/B))$ proportional fairness for $N$ agents, as long as the predictions are reasonably accurate (constant multiplicative error). We also show this to be optimal given predictions.

Finally, in Section 5 we extend our model even further by allowing a batch of $L$ public goods to arrive in each round $t$ and achieve $O(\log(\min(N, L) \cdot T/B))$ proportional fairness. In fact, we show that this model strictly generalizes not only our initial public-goods model, but also the private-goods model of Banerjee et al. [2022a], and our positive result in this very broad model (almost) implies theirs.

### 1.2 Related Work

#### Allocation of private goods.

The majority of the prior work on online fair division has focused on private goods, for which achieving even basic notions of fairness comes at the cost of extreme inefficiency in the absence of any predictions regarding agents’ values for future goods [Benade et al., 2018; Zeng and Psomas, 2020]. Banerjee et al. [2022a] show that total value predictions can be leveraged to achieve improved fairness guarantees with respect to the Nash social welfare objective. Note that assuming access to predictions of agents’ total values for all goods is related to work which assumes that the values of each agent are normalized to add up to 1 [Gkatzelis et al., 2021; Barman et al., 2022] or that they are drawn randomly from a normalized distribution [Bogomolnaia et al., 2019].

Our techniques are motivated by the set-aside greedy algorithm in [Banerjee et al., 2022a] and we generalize their finding that predictions help improve fairness guarantees to a more general setting with public goods. This introduces new challenges which we briefly touch on below.

#### Algorithm design.

Banerjee et al. [2022a] act on the private good in each round by allocating half of it equally among the agents and the other half greedily to maximize the predicted Nash welfare. In our model, the overall budget $B$ crucially connects the different rounds. A natural extension of their approach would be to decouple the rounds by setting an “artificial” per-round budget of $B/T$, but it is easy to see that this would incur $\Omega(T/B)$ proportional fairness, which is much worse than the $O(\log(T/B))$ approximation we obtain.

To achieve our optimal approximation guarantee for public goods, we instead need to: 1) Implement the “set-aside step” in a more careful way (see Algorithm 1), and 2) Dynamically control the amount of budget used by the “greedy step” by adding a linear penalty to the optimization problem solved in each stage. The resulting algorithm is quite different: whereas the previous algorithm always spends a budget of $\frac{B}{N}$ in each round on the greedy allocation, our algorithm may spend as little as 0 (in rounds where the values are low), as much as $1 - \frac{B}{T}$ (in rounds where the values are high), or any amount in between.

#### Lower bounds.

Our lower bounds (Theorem 1 and 4) are novel and non-trivial. The worst-case instances we design have a very different structure than the instances used by Banerjee et al. [2022a] because in our instances some of the agents “cooperate” (i.e., they want the algorithm to allocate to the common public goods they like), which is an aspect missing in the context of private goods. We also point out that their lower bounds are slightly loose ($\Omega(\log^{1-c} N)$ and $\Omega(\log^{1-\epsilon} T)$) whereas ours are asymptotically tight ($\Omega(\log N)$ and $\Omega(\log(T/B))$).

#### Allocation of public goods.

Much of the literature on public goods focuses on the offline setting – agents have ap-
proval preferences, and we have a fixed budget (i.e., the offline version of the setting in Section 3). This offline version has been studied under various names, such as probabilistic voting [Bogomolnaia et al., 2005], fair mixing [Aziz et al., 2019], fair sharing [Duddy, 2015], portioning [Brandl et al., 2021], and randomized single-winner voting [Ebadian et al., 2022b]. The latter shows that with access to only ranked preferences, $\Theta(\log T)$ is the best possible proportional fairness in the offline setting. Interestingly, we show that incomplete information resulting from online arrivals leads to the same $\Theta(\log T)$ proportional fairness.

Multi-winner voting extends this by selecting a subset of $k$ candidates, which is the offline version of our model in Section 4 with $B = k$. Multi-winner voting can be extended to fair public decision-making [Conitzer et al., 2017; Garg et al., 2021] and participatory budgeting [Aziz and Shah, 2021; Fain et al., 2016], which are further generalized by the public good model of Fain et al. [2018]. These generalizations allow more complex feasibility restrictions on the allocations than our budget constraint, but work in the offline setting. Freeman et al. [2017] consider optimizing the Nash welfare in a model similar to ours. However, they do not provide any approximation guarantees; instead, they study natural online algorithms from an axiomatic viewpoint. The work of Do et al. [2022] also studies the allocation of indivisible public goods that arrive online, but from an axiomatic approach or by making distributional assumptions. Oren and Lucier [2014] also consider the allocation of indivisible public goods, but in the setting where the agents arrive online rather than the goods. Lastly, Lackner [2020] studies voting over rounds from an axiomatic approach for the indivisible case.

Primal-dual analysis. Our competitive ratios are derived using primal-dual analyses (see [Buchbinder and Naor, 2009] for an excellent overview). Almost all of this work deals with additive objective functions. Two notable exceptions to this are the work of Devanur and Jain [2012], who show how to extend these approaches to non-linear functions of additive rewards, as well as Azar et al. [2016], who consider a variant of the proportional allocation objective, which require additional boundedness conditions on the valuations. Barnas et al. [2020] show how to adapt primal-dual analyses to incorporate predictions for several single-agent settings.

2 Model

We study an online resource allocation problem with $N$ agents and $T$ rounds. Our algorithms observe $N$, but they might not know the number of rounds $T$, in advance. We study both horizon-aware algorithms which know $T$ and horizon-independent algorithms which do not. For simplicity we use $[k]$ to denote the set $\{1, \ldots, k\}$ for a given $k \in \mathbb{N}$.

Online arrivals. In the basic version of the model, in each round $t \in [T]$, a public good, which we refer to as good $t$, "arrives". Upon its arrival, we learn the value $v_{i,t} \geq 0$ of every agent $i \in [N]$ for it. In Section 5, we extend the model to allow a batch of $L$ public goods arriving in each round.

Online allocations. When good $t$ arrives, the online algorithm must irrevocably decide the allocation $x_t \in [0, 1]$ to good $t$, before the next good arrives.\footnote{In this work we focus on deterministic algorithms, but this is w.l.o.g. since we consider fractional allocations.} We use $x = (x_t)_{t \in [T]}$ to denote the final allocation computed by the online algorithm. In the absence of any further constraints, the decision would be simple: allocate as much as possible to every good by setting $x_t = 1$ for each $t \in [T]$. However, the extent to which the algorithm can allocate these goods is limited by an overall budget constraint: $\sum_{t=1}^{T} x_t \leq B$, where $B \geq 0$ is a fixed budget known to the online algorithm in advance.

Linear agent utilities. Choosing to allocate $x_t$ to good $t$ simultaneously yields utility $v_{i,t} \cdot x_t$ to every agent $i \in [N]$. Moreover, we assume that agent utilities are additive across goods, i.e., the final utility of agent $i$ is given by $u_i(x) = \sum_{t=1}^{T} v_{i,t} \cdot x_t$. This class of linear agent utilities is widely studied and it admits several natural interpretations, depending on the application of interest. In applications like budget division, each public good $t$ is a project (e.g., an infrastructure project), and $x_t$ is the amount of a resource (e.g., time or money) invested in the project. In applications such as participatory budgeting or voting, each public good $t$ is an alternative or a candidate, and $x_t$ is the (marginal) probability of it being selected (indeed, one can compute a lottery over subsets of goods of size at most $B$ under which the marginal probability of selecting each good $t$ is precisely $x_t$).

When working with fractions $x/y$ with $x, y \geq 0$, we adopt the convention that $x/y = 1$ when both $x = y = 0$, while $x/y = +\infty$ if $y = 0$ and $x > 0$. We use $H_k = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{k}$ to denote the $k$-th harmonic number.

2.1 Proportional Fairness

We want the allocation $x$ computed by our online algorithm to be fair. In this work, we use the notion of proportional fairness, which is a quantitative fairness notion that was first proposed in the context of rate control in communication networks [Kelly et al., 1998].

Definition 1 (Proportional Fairness). For $\alpha \geq 1$, allocation $x$ is called $\alpha$-proportionally fair if for every other allocation $w$ we have $\frac{1}{N} \sum_{i=1}^{N} \frac{u_i(w)}{u_i(x)} \leq \alpha$. If $x$ is $1$-proportionally fair, we simply refer to it as proportionally fair\footnote{The $1$-proportional fair criterion is more commonly (but equivalently) written as $\frac{1}{N} \sum_{i=1}^{N} \frac{u_i(w)}{u_i(x)} \leq 1$.}. We say that an online algorithm is $\alpha$-proportionally fair if it always produces an $\alpha$-proportionally fair allocation.

It is known that in the offline setting, where all agent values are known up front, a $1$-proportionally fair allocation $x$ always exists, and this is the lowest possible value of proportional fairness [Fain et al., 2018]. It is also known that proportional fairness is a strong guarantee that implies several other guarantees sought in the literature. Below, we show two examples: the core and Nash social welfare.

Proportional fairness implies the core. For $\alpha \geq 1$, allocation $x$ is said to be in the $\alpha$-core if there is no subset of agents $S$ and allocation $w$ such that $\sum_{i \in S} u_i(w) \geq \alpha \cdot u_i(x)$ for all $i \in S$ and at least one of these inequalities is strict. We say that an online algorithm is $\alpha$-core if it always produces
an allocation in the $\alpha$-core. The following is a well-known relation between proportional fairness and the core.

**Proposition 1.** For $\alpha \geq 1$, every $\alpha$-proportionally fair allocation is in the $\alpha$-core.

**Proportional fairness implies optimal Nash welfare.** A common objective function studied in multi-agent systems is the Nash social welfare, which aggregates individual agent utilities into a collective measure by taking the geometric mean. That is, the Nash social welfare of allocation $x$ is given by $\text{NSW}(x) = \left( \prod_{i=1}^{N} u_i(x) \right)^{1/N}$. For $\alpha \geq 1$, we say that allocation $x$ achieves an $\alpha$-approximation of the Nash welfare if $\text{NSW}(x) \leq \alpha$ for all allocations $w$. We say that an online algorithm achieves an $\alpha$-approximation of the Nash welfare if it always produces an allocation that achieves an $\alpha$-approximation of the Nash welfare. It is also well-known that $\alpha$-proportional fairness implies an $\alpha$-approximation of the Nash welfare (in particular, a proportionally fair allocation has the maximum possible Nash welfare).

**Proposition 2.** For $\alpha \geq 1$, if allocation $x$ is $\alpha$-proportionally fair, then $x$ achieves an $\alpha$-approximation of the Nash welfare.

We remark that the upper bounds derived in this work hold for the stronger notion of proportional fairness, while the lower bounds hold even for the weaker notion of Nash welfare approximation.

### 2.2 Set-Aside Greedy Algorithms

In order to compute (approximately) proportionally fair allocations, we consider a family of online algorithms, called *Set-Aside Greedy Algorithms*. Recent work [Banerjee et al., 2022a; Barman et al., 2022] has demonstrated how such algorithms can be used to get strong performance guarantees for the online allocation of private goods; we show that with considerable modifications, they can also achieve compelling fairness guarantees for allocating public goods.

At a high level, a set-aside greedy algorithm divides the overall budget $B$ into two equal portions.

1. The first half, called the *set-aside budget*, is used to allocate $y_t \in [0, 1]$ to each good $t$ in such a manner that $\sum_{t=1}^{T} y_t \leq B/2$ and this portion of the allocation guarantees each agent $i$ a certain minimum utility of $\Delta_i$ (i.e., $\sum_{t=1}^{T} v_{i,t} \cdot y_t \geq \Delta_i$). For example, if $y_t = B/(2T)$ for each $t \in [T]$, then we can use $\Delta_i = B/(2T) \cdot \sum_{t=1}^{T} v_{i,t}$. This ensures that in the proportional fairness definition (Definition 1), the ratio $\frac{u_i(x)}{u_i(w)}$ is not excessively large for any agent $i$.

2. The second half, called the *greedy budget*, is used to allocate $z_t \in [0, 1 - y_t]$ to each good $t$ in such a manner that $\sum_{t=1}^{T} z_t \leq B/2$. This portion of the budget is used in a adaptive greedy-like fashion toward online optimization of the desired objective.

We refer to $y_t$ and $z_t$ as *semi-allocations* to good $t$, and the final allocation of good $t$ is determined by combining these two semi-allocations, i.e., $x_t = y_t + z_t$. An important quantity in both our algorithm and its analysis is the promised utility to an agent, defined as follows.

**Definition 2 (Promised Utility).** The semi-allocations $y_1, \ldots, y_T$ guarantee that by the end of the algorithm each agent $i$ will receive a utility of $\sum_{t=1}^{T} v_{i,t} \cdot y_t \geq \Delta_i$ from the set-aside portion of the budget. By round $t$ the algorithm has already set semi-allocations $z_1, \ldots, z_{t-1}$ using the greedy portion of the budget, and needs to now decide $z_t$. At this stage, as a function of $z_t$, the algorithm can guarantee that each agent $i$ will eventually receive utility at least $\bar{u}_{i,t}(z_t) = \Delta_i + \sum_{\tau=1}^{t-1} v_{i,\tau} \cdot z_{\tau}$, even if they do not benefit from any more of the greedy budget. We refer to this as the promised utility.

### 3 Warm Up: Binary Utilities and Unit Budget

Before presenting our main results, we first build some intuition regarding our online setting and the proportional fairness objective by considering the interesting special case where agents have binary utilities for goods (i.e., $v_{i,t} \in \{0, 1\}$ for each $i, t$) and the total budget is $B = 1$. In this setting, which is motivated by approval voting, we say that agent $i$ “likes” good $t$ if $v_{i,t} = 1$, and does not like $t$ otherwise. Note that with $B = 1$, the budget constraint is $\sum_{t=1}^{T} x_t \leq 1$, which means $x_t$ can be interpreted as the fraction of an available resource (e.g., time or money) that is dedicated to good $t$.

First, in the trivial case with a single agent ($N = 1$), we can simply set $x_t = 1$ when the first good $t$ liked by the agent arrives, which easily yields (exact) proportional fairness.

It is tempting to extend this idea to the case of $N > 1$ agents. However, we find that even in this restricted scenario with binary utilities and unit budget, no online algorithm achieves $o(\log N)$-proportional fairness, or even the weaker guarantee of $O(\log N)$-approximation of the Nash welfare. In fact, this remains true even if the algorithm is horizon-aware (i.e., knows $T$ in advance) and knows precisely how many goods each agent will like in total. Intuitively, this is because we show that no online algorithm can sufficiently distinguish between instances where many agents like the same goods and those where agents like mostly disjoint goods. We defer the proof to our full version [Banerjee et al., 2022b].

**Theorem 1.** With binary agent utilities ($v_{i,t} \in \{0, 1\}$, $\forall i, t$) and unit budget ($B = 1$), every online algorithm is $\Omega(\log N)$-proportionally fair (in fact, achieves $\Omega(\log N)$-approximation of the Nash welfare), even if the algorithm is horizon-aware and knows in advance the total number of goods each agent will like.

Next, we provide a set-aside greedy algorithm that achieves $O(\log N)$ proportional fairness (and therefore, $O(\log N)$-NSW optimality), thus establishing $\Theta(\log N)$ as the best possible approximation in this restricted scenario. We remark that this restricted case of binary utilities and unit budget already poses an interesting challenge since a constant approximation ratio is not possible. However, an $O(\log N)$ approximation is still quite reasonable as it does not depend on the horizon $T$ (which can typically be very large) and in practice the number of agents $N$ is reasonably small. We also remark that we achieve the $O(\log N)$ upper bound using a horizon-independent algorithm, while the lower bound of Theorem 1 holds even when the algorithm is horizon-aware.
Our algorithm, SETASIDEGREEDY-BINARY, works as follows: it uses the set-aside portion of the budget to set $y_t = 1/(2N)$ whenever good $i$ is the first liked good of at least one agent (note that $\sum_i y_t \leq 1/2$). This ensures that each agent $i$ gets utility at least $\Delta_i = 1/(2N)$. Based on this, the algorithm uses the following expression of promised utility to agent $i$ in round $t$: $\tilde{u}_{i,t}(z_t) = \frac{1}{2N} + \sum_{t'=1}^t v_{i,t'} \cdot z_{t'}$. On the other hand, $z_t$ is chosen to be the smallest possible value such that, for each agent $i$, the ratio of her value $v_{i,t}$ for good $t$ to her promised utility $\tilde{u}_{i,t}(z_t)$ is at most a target quantity. For a formal outline, see our full version [Banerjee et al., 2022b].

The theorem below shows that algorithm SETASIDEGREEDY-BINARY is $O(\log N)$ proportionally fair for binary utilities and unit budget. We defer the proof to our full version [Banerjee et al., 2022b].

**Theorem 2.** Algorithm SETASIDEGREEDY-BINARY with $\alpha \geq 2 \ln(2N)$ realizes an $\alpha$-proportional fair allocation.

## 4 General Utilities and Budget

Having built some intuition about the online setting and the proportional fairness objective by considering the restricted special case of the problem wherein all values $v_{i,t}$ are in $\{0, 1\}$ and the budget is $B = 1$, we now turn to the more general model described in Section 2. Recall that this model generalizes the setting in Section 3 in two ways:

1. Agent values $v_{i,t}$ can be any (non-negative) real number.
2. The budget constraint is $\sum_{t=1}^T x_t \leq B$, so that the per-round constraint of $x_t \leq 1$, for each $t \in [T]$, is no longer redundant.

### 4.1 The Case for Predictions

In this general case, we prove that the problem becomes significantly more difficult: without access to any predictions, every online algorithm is $\Omega(T/B)$-proportionally fair (in fact, achieves $\Omega(T/B)$-approximation of the Nash welfare), in stark contrast to the $O(\log N)$-proportional fairness guarantee in the previous section. The proof is on our full version [Banerjee et al., 2022b].

**Proposition 3.** Under general agent values and budget $B$, every online algorithm is $\Omega(T/B)$-proportionally fair (in fact, achieves $\Omega(T/B)$-approximation of the Nash welfare).

The hardness instance used above is specifically engineered to exploit the fact that the algorithm has no information about the input. However, in most practical settings, it is reasonable to assume that the algorithm has access to some information about the input. This could come from historical data, stochastic assumptions, or simply from properties of the problem at hand (e.g., if $v_{i,t}$ represents the monetary value that agent $i$ has for good $t$, then we may have bounds on how large $v_{i,t}$ can be.)

Motivated by this, we turn to the growing literature on prediction-augmented algorithms, and allow the algorithm access to additional side-information about agents’ valuations. Clearly, if all values $(v_{i,t})_{i \in [N], t \in [T]}$ are available beforehand, then the problem is trivial; the challenge lies in understanding what minimal additional information (or ‘prediction’) can lead to sharp improvements in performance, and how robust these improvements are to errors in these predictions. To this end, we now adapt an idea introduced by [Banerjee et al., 2022a] for online allocation with private goods, and assume that the algorithm has side information about each agent’s total value for all the goods.

**Definition 3 (Total Value Predictions).** For any agent $i$, we define her total value to be $V_i = \sum_{t=1}^T v_{i,t}$. For $c_i, d_i \geq 1$, $V_i$ is said to be a $[c_i, d_i]$-prediction of $V_i$ if $V_i = \left[\frac{1}{d_i}V_i, c_iV_i\right]$. In other words, $c_i$ and $d_i$ denote the multiplicative factors by which the prediction $V_i$ may overestimate and underestimate the value of $V_i$. When $c_i = d_i = 1$, we call them perfect predictions. In the next section, we assume that we have access to $V_i$ for each agent $i$, with $c_i$ and $d_i$ parameterizing the robustness of our algorithm, i.e., the degradation in its performance due to worsening predictions. Our algorithm does not need to know (an upper bound on) the $d_i$’s for tuning one of its parameters; it does not however need to know the $c_i$’s (these are only used in the analysis).

### 4.2 Proportional Fairness with Predictions

Using the above notion of predictions, we design algorithm SETASIDEGREEDY-GENERAL, a variant of our earlier Set-Aside Greedy algorithm that has a dramatically better proportional fairness guarantee compared to the hardness result of $\Omega(T/B)$ in Proposition 3. Given perfect predictions, our algorithm achieves a proportional fairness of $O(\log(T/B))$. Moreover, SETASIDEGREEDY-GENERAL turns out to be remarkably robust to prediction errors; for example, all our asymptotic guarantees remain unchanged as long as all the $c_i = O(1)$ and $d_i = O(\log(T/B))$.

We formally specify our SETASIDEGREEDY-GENERAL in our full version [Banerjee et al., 2022b], but at a high-level, it works as follows. As before, the algorithm splits the budget into two parts, and the total allocation of the good in round $t$ is obtained by adding the contributions (semi-allocations) from each part, i.e., $x_t = y_t + z_t$. The semi-allocation from the first (set-aside) part is $y_t = B/(2T)$ for each $t \in [T]$. This portion of the allocation guarantees each agent utility at least $\Delta_i = \frac{B}{2T} \cdot V_i$. Next, in each round $t \in [T]$, the algorithm uses the second part of the budget to compute a greedy semi-allocation $z_t$. This is done by choosing $z_t$ to optimize a function of the agents’ predicted promised utilities.

**Definition 4 (Predicted Promised Utility).** Given a prediction $V_i$ of the total value of agent $i$, the predicted promised utility of agent $i$ in round $t \in [T]$ is defined as $\tilde{u}_{i,t}(z_t) = \frac{B}{2T} \cdot V_i + \sum_{t'=1}^t v_{i,t'} \cdot z_{t'}$. We omit $z_1, \ldots, z_{t-1}$ from the argument of $\tilde{u}_{i,t}$ since they are fixed prior to round $t$.

Note that this quantity can be computed by the algorithm, as a function of $z_t$ it wants to choose, since it has knowledge of $V_i$ (prediction) and semi-allocations $(z_{t'})_{t' < t}$ from the previous rounds. We use these predicted promised utilities to achieve the following guarantee.

**Theorem 3.** For any $\alpha \geq 4 \ln \left(\frac{B}{2T}\right) + \frac{2}{N} \sum_i \ln(d_i)$, algorithm SETASIDEGREEDY-GENERAL produces a feasible allocation $x$, which satisfies $\max_{w} \frac{1}{N} \sum_{i=1}^{N} u_{i,w} \left(\frac{x_{i,w}}{x_i}\right) \leq \alpha$, where the maximum is over all feasible allocations $x$. 

Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-23)

24
This result is almost subsumed by our positive result (Theorem 5) for the more general model in the next section, in which we allow a batch of \( L \) public goods to arrive in each round; setting \( L = 1 \) recovers the bounds derived in this section. The only difference is that in the algorithm for the batched model (Algorithm 1), the greedy semi-allocation step is a convex optimization problem, which can be solved up to an \( \epsilon \) error; for the single-good-per-round case, we are able to replace this with a combinatorial step in SETASIDEGREEDY-GENERAL that can be performed exactly in polynomial time. This requires us to show that the greedy semi-allocations computed by this combinatorial step satisfy the same properties that a solution to the convex optimization problem would have satisfied. We defer this discussion to our full version [Banerjee et al., 2022b].

Let us consider the implications of Theorem 3. The expression in the statement of Theorem 3 is not exactly the proportional fairness objective, since the term for each agent \( i \) is scaled with a (potentially different) factor \( c_i \). Applying the arguments in Section 2.1, we can turn this into an approximation of proportional fairness, the core, and the Nash social welfare. The proof is in the supplementary material.

**Corollary 1.** For any \( \alpha \geq 4 \ln \left( \frac{4^k}{T} \right) + \frac{4}{T} \sum_i \ln(d_i) \), the allocation realized by SETASIDEGREEDY-GENERAL is

1. \( \alpha \cdot \max_{i \in [N]} c_i \)-proportionally fair, and hence in the \( \alpha \cdot \max_{i \in [N]} c_i \)-core
2. \( \alpha \cdot \prod_{i \in [N]} c_i \) - approximation of the Nash welfare.

### 4.3 Hardness with Predictions

Theorem 3 shows that in online allocation of public goods with general values and budget, having access to reasonable predictions of each agent’s total value can lead to a dramatic improvement in the proportional fairness guarantee from \( \Omega(T/B) \) to \( O(\log(T/B)) \). Given the size of the side information (which lies in \( \mathbb{R}^N \), since we need one prediction per agent) relative to the ambient size of the input (which lies in \( \mathbb{R}^{NT} \), with one valuation per agent per round), this is a surprising improvement in performance.

One may wonder whether these predictions are so strong that one can do even better. The following result shows, however, that even with a single agent, and perfect knowledge of her total value \( V_1 \), any online algorithm is \( \Omega(\log(T/B)) \)-proportionally fair. Thus SETASIDEGREEDY-GENERAL is essentially optimal for our setting. The proof is in our full version [Banerjee et al., 2022b].

**Theorem 4.** For \( N = 1 \) agent, every online algorithm is \( \Omega(\log(T/B)) \)-proportionally fair (in fact, achieves an \( \Omega(\log(T/B)) \)-approximation for the Nash welfare), even with perfect knowledge of horizon \( T \) and the total value of the agent \( V_1 = V_1 = \sum_{t=1}^{T} v_{i,t} \).

### 5 Extension to Batched Arrivals

In this section, we present our most general setting, which we refer to as the batched public goods model. This not only generalizes the single good per round model of Section 2, but also the setting of [Banerjee et al., 2022a] with private goods. In each round \( t \in [T] \), a batch of \( L \) public goods arrive (as opposed to a single public good). Upon their arrival, the algorithm learns the value \( v_{i,t} \) of each agent \( i \in [N] \) for each good \( l \in [L] \) in the batch. It then must irrevocably decide the allocation \( x_{i,t} \in [0,1] \) to each good \( l \) in the batch, before the next round. We use \( x_i = (x_{i,t})_{t \in [T]} \) to denote the allocation in round \( t \), and \( x = (x_t)_{t \in [T]} \) to denote the final allocation. We also incorporate two types of constraints on the allocation \( x \):

1. (Per-round constraint) \( \sum_{t=1}^{T} x_{i,t} \leq 1 \) for all \( t \in [T] \).
2. (Overall constraint) \( \sum_{t=1}^{T} \sum_{l=1}^{L} x_{i,t} \leq B \) for \( B \geq 0 \) known to the algorithm in advance.

The allocation \( x_{i,t} \) to good \( l \) in round \( t \) yields utility \( v_{i,t} \cdot x_{i,t} \) to every agent \( i \). We assume that agent utilities are additive, i.e., the final utility of agent \( i \) is given by \( u_i(x) = \sum_{t=1}^{T} \sum_{l=1}^{L} v_{i,t} \cdot x_{i,t} \). Our aim, as before, is to realize an \( \alpha \)-proportionally fair allocation (Definition 1 does not require any modifications except using these new utility functions) for the smallest possible \( \alpha \).

Note that the overall constraint is the same as in the model in Section 2 with a single good per round. However, in the batched model, the per-round constraint can place additional restriction on how much budget can be spent in any single round. Note also that choosing the per-round bound to be 1 is without loss of generality; in particular, since agent utilities are linear, having a per-round constraint of \( b \) can be reduced to our setting by scaling each allocation, as well as the total budget, by a factor of \( b \). Finally, the per-round constraint becomes vacuous if \( B \leq 1 \), and the overall constraint becomes vacuous if \( B \geq T \); therefore, we assume, without loss of generality, that \( 1 \leq B \leq T \).

This model captures the following special cases. Before we dive into the algorithm and analysis, we briefly mention some special cases of interest which this model generalizes.

1. **Single public good.** When \( L = 1 \), we trivially recover the setting of Section 4, where there is one public good in each round.
2. **Single private good.** When \( L = N, B = T \), and \( v_{i,t} = 0 \) if \( i \neq j \), we recover the setting studied by [Banerjee et al., 2022a]. In their setting, there is a single private good arriving in each round, which the algorithm needs to split among the \( N \) agents. When cast in our model, \( v_{i,t} \) is the value that agent \( i \) has for the good in round \( t \), and \( x_{i,t} \) is the fraction of good \( t \) that agent \( i \) is allocated. They only study per-round constraints, so one of our contributions is a generalization of their result to the budgeted setting.
3. **Batched private goods.** When \( L = L \cdot N, B = T \), and \( v_{i,t} = 0 \) if \( i \neq j \mod N \), we capture a setting where there are \( L \) private goods arriving in each round, and the algorithm can (fractionally) allocate at most 1 good in total among the agents in each round.

#### 5.1 The Set-Aside Greedy Algorithm for Batched Public Goods

We present Algorithm 1 for the batched public goods model which generalizes our guarantees from Section 4 and, partly,
Algorithm 1 Set-Aside Greedy Algorithm for Batched Public Goods

**Input:** Target proportional-fairness level \( \alpha \); total value predictions \((\tilde{V}_i)_{i \in [N]}\)

1. Define \( q = \frac{B}{2\min(N,L)} \).
2. for all \( t = 1 \) to \( T \) do
3.   Given values \((v_{i,l,t})_{i \in [N], t \in [L]}\); find ‘favorite goods’ set \( F_t \) as follows: Initialize \( F_t \leftarrow \emptyset \).
4.   for all \( l = 1 \) to \( N \) do
5.     Let \( l_t \leftarrow \arg \max_i \{v_{i,l,t}\} \) (breaking ties arbitrarily), and update \( F_t \leftarrow F_t \cup \{l_t\} \).
6.   end for
7.   Set-aside part. Set \( y_{l,t} = \frac{B}{2\min(N,L)} \cdot \mathbb{1}\{l \in F_t\} \).
8.   Greedy part. Let \((z_t, \lambda_t)\) be an optimal solution to the following optimization problem:
   \[
   \text{Maximize} \quad \frac{1}{N} \sum_{i=1}^{N} \ln \left( \tilde{u}_{i,t}(z) \right) + \lambda q \text{ subject to } \sum_{l=1}^{L} z_l + \lambda = 1 - \frac{B}{2\min(N,L)} \text{ and } z, \lambda \geq 0. 
   \]
9.   Allocate \( x_{l,t} = \tilde{y}_{l,t} + z_{l,t} \).
10. end for

Theorem 5. In the batched public goods model, for any \( \alpha \geq 4 \ln \left( \frac{2\min(N,L)T}{B} \right) + \frac{4}{N} \sum_i (d_i) \), Algorithm 1 produces a feasible allocation \( x \), which satisfies \( \max_w \frac{1}{N} \sum_{i=1}^{N} \frac{u_i(w)}{w} \leq \alpha \), where the maximum is over all feasible allocations \( w \).

Using exactly the same proof as that of Corollary 1, we obtain the following guarantees for Algorithm 1 with respect to proportional fairness, the core, and the Nash social welfare.

**Corollary 2.** For \( \alpha \geq 4 \ln \left( \frac{2\min(N,L)T}{B} \right) + \frac{4}{N} \sum_i (d_i) \), Algorithm 1 is
1. \((\alpha \cdot \max_i c_i)\)-proportionally fair, and hence in the \((\alpha \cdot \max_i c_i)\)-core.
2. achieves \((\alpha \cdot \prod_{i \in [N]} c_i)^{\frac{1}{N}}\)-approximation of the Nash social welfare.

Recall that the private goods setting of Banerjee et al. [2022a] is a special case of our setting in which \( L = N \), \( B = T \), and the valuation matrix of the \( N \) agents for the \( N \) public goods in each round is a diagonal matrix. For this special case, our Nash welfare approximation is \( O((\prod_{i \in [N]} c_i)^{1/N} \cdot (\ln N + (1/N) \cdot \sum_i (d_i)) \). The Nash welfare approximation obtained by Banerjee et al. [2022a] is almost the same, except that \( \ln N \) is replaced by \( \ln (\ln N, \ln T) \). Thus, for \( T = \Theta(N) \), our result strictly generalizes theirs and they prove this bound to be almost tight. For \( T = o(N) \), they derive a better approximation that depends on \( \ln T \) instead of \( \ln N \), (and it is unknown if this is tight). It would be interesting to see if our result can also be improved in this case.

6 Discussion and Future Directions

While we focus on the proportional fairness objective, a natural question is whether our results can be extended to other non-linear objective functions, such as generalized p-mean welfare measures [Barman et al., 2022; Ebadian et al., 2022a]. An other interesting direction for future research is to develop a better understanding of appropriate types of predictions for online fair division. For example, what if the algorithm is also provided with a prediction regarding the total value of each good across all agents, but not specifying which of the agents will like it and by how much? A related question to the one above is in regards to the process for generating agent valuations. In our model, the values \( V = (v_{i,l})_{i \in [N], t \in [T]} \) are allowed to arrive in an adversarial order. What if we consider slightly more optimistic models such as the random-order model or the stochastic model [Borodin and El-Yaniv, 2005]?

Acknowledgements

We are grateful for financial support from NSF CAREER award CCF 2047907 (VG), an NSERC Discovery Grant (EM and NS), an NSERC fellowship PGSD3-532673-2019 (BJ), and AFOSR grant FA9550-23-1-0068, NSF grant ECCS-1847393, and the Simons Institute for the Theory of Computing (BJ and SB).
References


