Allocation for Social Good: Auditing Mechanisms for Utility Maximization

TAYLOR LUNDY, University of British Columbia

ALEXANDER WEI, Harvard University

HU FU, University of British Columbia

SCOTT DUKE KOMINERS, Harvard University

KEVIN LEYTON-BROWN, University of British Columbia

We consider the problem of a nonprofit organization (“center”) that must divide resources among subsidiaries (“agents”), based on agents’ reported demand forecasts, with the aim of maximizing social good (agents’ valuations for the allocation minus any payments that are imposed on them). We investigate the impact of a common feature of the nonprofit setting: the center’s ability to *audit* agents who receive allocations, comparing their actual consumption with their reported forecasts. We show that auditing increases the power of mechanisms for utility maximization, both in unit-demand settings and beyond: in unit-demand settings, we consider both constraining ourselves to an allocation function studied in past work and allowing the allocation function to vary; beyond unit demand, we adopt the VCG allocation but modify the payment rule. Our ultimate goal is to show how to leverage auditing mechanisms to maximize utility in repeated allocation problems where payments are not possible; we show how any static auditing mechanism can be transformed to operate in such a setting, using the threat of reduced future allocations in place of monetary payments.

Additional Key Words and Phrases: market design; AI for social good; fair resource allocation; utility maximization; repeated games

ACM Reference Format:

Taylor Lundy, Alexander Wei, Hu Fu, Scott Duke Kominers, and Kevin Leyton-Brown. 2019. Allocation for Social Good: Auditing Mechanisms for Utility Maximization. In *ACM EC '19: ACM Conference on Economics and Computation (EC '19), June 24–28, 2019, Phoenix, AZ, USA*. ACM, New York, NY, USA, 19 pages. <https://doi.org/10.1145/3328526.3329623>

We gratefully acknowledge support from our funders: National Science Foundation grants SciSIP-1535813 and SES-1459912, the Ng Fund and the Mathematics in Economics Research Fund of the Harvard Center of Mathematical Sciences and Applications, and the Human Capital and Economic Opportunity Working Group sponsored by the Institute for New Economic Thinking (Kominers); the NSERC Discovery Grants program (Fu; Leyton-Brown); DND/NSERC Discovery Grant Supplement, Facebook Research, Canada CIFAR AI Chair Amii (Leyton-Brown), Canada Research Chair, NSEERC Discovery Grant Accelerator (Fu).

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

EC '19, June 24–28, 2019, Phoenix, AZ, USA

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6792-9/19/06...\$15.00

<https://doi.org/10.1145/3328526.3329623>

1 INTRODUCTION

Foundations, governments, and large charitable organizations are often responsible for dividing scarce resources among smaller organizations or subsidiaries in a way that maximizes the social good achieved. For example, area food banks address regional food shortages by distributing food to local food pantries and soup kitchens. Aid organizations like the International Red Cross, meanwhile, allocate supplies to local disaster relief partners in hopes of getting those supplies to end recipients as quickly and effectively as possible. And funding agencies like the NSF distribute resources across research institutions with an eye towards maximizing valuable output.

Resource allocation in the settings just described is challenging because (1) the smaller subsidiaries (the “agents”) have access to private information about local demand that the large organization (the “center”) needs to learn in order to decide upon an allocation; yet (2) while both the center and the agents care about the overall welfare, the agents are at least somewhat more motivated to help their own clients than those of other agents. The center thus has reason to worry that the agents are not incentivized to report their private information completely honestly. We consider solutions to the center’s problem that draw on the formal tools of mechanism design.

The classical mechanism design approach to incentivizing truthful reporting is to allocate resources via an appropriately designed auction. Indeed, the literature has much to say about how to do so: the VCG mechanism offers agents dominant strategies for truthfully reporting their private information and—if they do so—guarantees efficient allocations. Unfortunately, however, VCG and its cousins target the wrong objective function for our setting: they aim to maximize surplus without consideration of the payments imposed—that is, they maximize allocative *efficiency*—whereas we actually want to maximize *total utility*, i.e., the overall social good including payments. We focus on total utility maximization for practical reasons: A food bank, for example, allocates food with the sole aim of improving social outcomes. Any money food pantries have to pay reduces their ability to staff their operations and/or to buy food in the open market. Moreover, unlike in other market design settings, food pantries facing higher demand have absolutely no increased ability to pay for food. Indeed, quite the opposite—the food pantries facing the highest demand are often especially under-resourced, either because they are located in lower-income areas, or because they are subject to unexpected food supply contractions. All food pantries give their food away; the crucial question for the center is whether a given food pantry will be able to make use of a food donation before it spoils.

The literature on mechanisms for utility maximization is relatively recent, starting with the independent work of Chakravarty and Kaplan [7] and Hartline and Roughgarden [19]; the mechanisms those authors derived naturally aim to minimize agents’ payments because of payments’ aversive impact on total utility.¹ However, in charitable contexts, the imposition of even small payments may be unpalatable—and moreover, as we have already described, ability to pay may be inversely correlated (or simply uncorrelated) with need (see, e.g., Dworczak & Kominers & Akbarpour [9]). We thus seek to further reduce or even eliminate utility maximizing mechanisms’ reliance on payments. Our main tool for doing so is another property of the charitable setting that does not always hold more generally: the ability (and often the legal requirement) for the center to *audit* agents to ensure that they make appropriate use of their allocations.

Of course, if we assume that the center can perfectly verify whether agents reported truthfully, then we (trivially) get around the need to impose payments. However, perfect auditing is unlikely to be available in practice. We thus assume instead that (1) agents face stochastic demand, about which they have only partial information *ex ante*, and (2) it is only possible to audit the allocations

¹ As is standard in mechanism design, payments are assumed to be collected by an outside third party, rather than somehow being integrated into welfare as a lump sum.

that agents actually receive—that is, the center cannot verify whether an unallocated agent would have been able to make use of an allocation. The key theme of our paper is that such auditing can lower requisite payments, and hence raise total utility.

We begin by introducing auditing into mechanisms for utility maximization in unit-demand settings. We show two results here: first, that we can improve existing mechanisms even if we are only able to augment the payment rule to incorporate auditing, leaving the allocation rule unchanged; second, we identify the optimal mechanism that is never allowed to pay the agents, which involves a new allocation rule.

We then expand our focus beyond the unit-demand setting. Little is known about mechanisms that maximize utility in more general settings, so we confine ourselves to the VCG allocation and ask whether auditing enables us to reduce the requisite payments. We obtain two positive results: first, a general-purpose modification of VCG that strongly resembles our mechanism for the unit-demand setting; and second, another VCG variant that further improves utilities under a distributional assumption (roughly, that agents' potential types admit an ordering by stochastic dominance).

Our ultimate goal is to derive a solution that works in settings in which we are unable to impose payments at all, but in which the allocation problem is indefinitely repeated (as it is, for example, in the food bank example); we show how to incentivize agents in such settings via the threat of reduced future allocations. Specifically, we show how to transform any static auditing mechanism *with payments* into a repeated, non-monetary auditing mechanism *without payments* that achieves the same expected utility across agents.

2 RELATED WORK

As mentioned in the Introduction, we build on existing work that studies mechanism design for utility maximization. The first major result and the most technically similar work in this area is the analysis of multi-unit “money-burning auctions” studied independently by Chakravarty and Kaplan [7] and Hartline and Roughgarden [19]. Those papers describe the Bayesian optimal social utility auction in the unit demand setting (without auditing) when no negative transfers are permitted. Our results make direct comparisons to the optimal mechanism Chakravarty and Kaplan [7] and Hartline and Roughgarden [19] describe; we show that auditing allows us to obtain higher-utility allocations. Similarly, some have sought to minimize payments in more general settings with higher-dimensional type spaces. One line of work aims to maintain efficiency but reduce the money collected in VCG auctions by redistributing payments back to the agents [4, 14, 15]. de Clippel et al. [8] investigate relaxing efficiency in order to further reduce payments. It is important to note that those papers relax the no negative transfers constraint to a budget balance constraint and therefore cannot be implemented via money burning in a setting without payments.

Similar to our own investigation of auditing, various other recent work focuses on using external signals to aid in mechanism design. Ben-Porath et al. [2], for example, find a class of incentive compatible mechanisms for a setting in which the center can verify agents' private information at a cost [see also 10, 11]. Other examples—albeit relatively far afield ones—include building auditing schemes that deter attackers in security games [3] and determining the subset of types that must be audited for an agent to have no benefit from lying [6].

Another branch of work considers a tool more general than ours: *contingent payments*, in which payments that are made conditional on an external signal. The external signal that is revealed to the mechanism can be any additional piece of information that may or may not have been known to the agent ahead of time. Auditing is the special case of contingent payments in which the signal that is revealed is an agents' *ex post* value for an allocation, which the agent was uncertain about prior to reporting. There are some general results that are known in contingent payment settings

relating to revenue maximization [12, 18]. There are also many examples of contingent payments being used in applications to both increase revenue and balance risk, including oil-lease auctions [17], ad auctions [26], and publishing rights [5]. The closest work in the contingent payment framework seeks to use contingent payments to increase social welfare. Most notably, Ma et al. [21] investigate maximizing both individual welfare and some global welfare function that depends on the consumption of an allocated item. Our mechanisms overlap with those of Ma et al. [21] in some special cases (discussed in the sequel); however, a key difference is that Ma et al. [21] do not focus on minimizing payments.

Finally, we draw on a literature on repeated allocation without money. A few papers have extended single round non-monetary mechanisms into truthful repeated mechanisms. For example, the scrip system of Gorokh et al. [13] allocates goods in a finite time-horizon model. Two additional papers study the time-discounted infinite horizon model. Guo et al. [16] described a repeated model that achieves incentive compatibility in exchange for relaxing efficiency; Balseiro et al. [1] proposed an incentive-compatible mechanism that approaches full social welfare extraction as the discount factor approaches 1. However, unlike our own work, both of those studies worked in settings with symmetric agents and a single good. Prendergast [23, 24], meanwhile, helped design a scrip currency system that the nonprofit Feeding America uses for equitable, repeated allocation of food to food banks (see also Kominers and Lam [20]).

3 PRELIMINARIES

We now formalize our model of utility maximizing mechanisms with auditing. In this model, a center distributes multiple units of a good to multiple agents, based on their reported demand forecasts. Payment is delayed until after the allocation and can depend on the outcome of an “audit,” which measures each agent’s realized consumption of any allocated goods. We then describe our solution concept and recall some auction theory results that will be useful in the rest of the paper.

3.1 Agents and Types

An *auditing game* consists of a single center and N agents. The center starts with M identical units of a good and wishes to distribute these units to the agents while maximizing total utility over all agents. Each agent has uncertain demand, characterized by a type t_i . We assume that agent i ’s type lies in a type space T_i and denote the product of all type spaces by T . We model the uncertainty in the demand of agent i as a non-negative random variable d_i with distribution given by the agent’s type t_i . That is, we identify each $t_i \in T_i$ with a distribution over the non-negative natural numbers and say that $d_i \sim t_i$. We assume realized demands d_i for each agent are drawn independently from their respective distributions. Thus, the agents’ types completely characterize the joint distribution of the realized demands. Finally, we let F_{t_i} denote the cumulative distribution function (CDF) of t_i , i.e., $F_{t_i}(k)$ is the probability that agent i has demand at most k .

We work with a Bayesian game setting. At the start of the game, the type t_i of agent i is drawn privately and independently according to some prior distribution G_i over T_i . We make the standard assumption that the type spaces T_i and the priors G_i are common knowledge. Thus, at the start of the game, each agent knows the prior distribution of each of the other agents’ types and their own realized type (which in turn defines a distribution over demand), while the center knows the prior distribution over each of the agents’ types.

We assume that agents have quasilinear utility: if agent i has realized demand $d_i \sim t_i$, receives an allocation of x_i units of the good, and is charged a payment of p_i , then agent i ’s utility is $\min(d_i, x_i) - p_i$. Observe that we thus assume that agents derive the same utility for each unit of the item they use and no utility from unused units, and that all agents obtain the same utility for consuming a unit of the good. We furthermore assume that agents are risk neutral.

3.2 Mechanism Design with Auditing

A (*quasilinear*) mechanism $\mathcal{M} = (\mathbf{x}, \mathbf{p})$ is defined by an allocation rule $\mathbf{x}$ and a payment rule $\mathbf{p}$, which are (possibly random) functions mapping N -tuples of types to N -tuples of natural numbers and reals respectively. The outputs of $\mathbf{x}$ and $\mathbf{p}$ represent allocations to and payments from each agent, respectively.

Definition 3.1. We say that $\mathcal{M} = (\mathbf{x}, \mathbf{p})$ is an *auditing mechanism* if it operates as follows (note that (1), (2), (3) and (6) are standard; our innovation is (4) and (5)).

- (1) Private types $t_i \in T_i$ are realized.
- (2) Each agent reports a type $\hat{t}_i \in T_i$ to the center; we use $\hat{\mathbf{t}} := (\hat{t}_1, \dots, \hat{t}_N)$ to denote the resulting reported type profile.
- (3) The center makes an allocation $\mathbf{x}(\hat{\mathbf{t}}) = (x_1(\hat{\mathbf{t}}), \dots, x_N(\hat{\mathbf{t}})) \in \mathbb{N}_{\geq 0}^n$, with $\sum_i x_i(\hat{\mathbf{t}}) \leq M$.
- (4) Each agent i 's demand $d_i \sim t_i$ is realized.
- (5) The center audits the agents and observes a level of consumption $d_i^{\text{obs}} := \min(d_i, x_i)$ for each agent $i \in \{1, \dots, N\}$.
- (6) The center charges a payment $p_i(\hat{\mathbf{t}}, d_i^{\text{obs}}) \in \mathbb{R}_{\geq 0}^n$ to agent i .

We say a mechanism is *non-audited* if its payment does not depend on the observed demand, in which case we may write $p_i(\hat{\mathbf{t}}, d_i^{\text{obs}})$ as $p_i(\hat{\mathbf{t}})$.

We observe that (6) restricts payments to be nonnegative (i.e., the mechanism never pays agents); this is a substantive assumption, as we discuss in the Conclusion. We insist on nonnegative payments because this allows us to identify auditing mechanisms that we can eventually transform into the repeated setting without money. In the transformed domain, payments from the agent to the center are replaced by “money burning”—the choice not to allocate certain goods to any agent. There is no clear analogous way to handle payments from the center to one or more agents—this would require somehow creating new copies of goods.

To define our solution concept, we introduce interim concepts of allocations, payments, and utilities. Specifically, let

$$x_i(\hat{t}_i) := \mathbf{E}_{t_{-i} \sim G_{-i}} [x_i(\hat{t}_i, \mathbf{t}_{-i})],$$

$$p_i(\hat{t}_i \mid d_i) := \mathbf{E}_{t_{-i} \sim G_{-i}} [p_i((\hat{t}_i, \mathbf{t}_{-i}), \min(d_i, x_i(\hat{t}_i, \mathbf{t}_{-i})))]$$

be the *interim allocation* and *interim payment* given the realized demand, respectively. The value $x_i(\hat{t}_i)$ represents the expected allocation received by agent i if they report $\hat{t}_i$ while all other agents report truthfully. The meaning of $p_i(\hat{t}_i \mid d_i)$ is similar, but with additional dependence on the realized demand d_i .

It will be convenient to have notation for agent i 's *interim utility* under a given (implicit) mechanism $\mathcal{M} = (\mathbf{x}, \mathbf{p})$:

$$u_i(\hat{t}_i, t_i) := \mathbf{E}_{d_i \sim t_i, \mathbf{t}_{-i} \sim G_{-i}} [\min(d_i, x_i(\hat{t}_i, \mathbf{t}_{-i})) - p_i(\hat{t}_i, d_i)].$$

We are now ready to define Bayes–Nash incentive compatibility.

Definition 3.2. An auditing mechanism $\mathcal{M}$ is *Bayes–Nash incentive compatible* (BIC) if it makes honest reporting a Bayes–Nash equilibrium, i.e., if under $\mathcal{M}$, we have $u_i(t_i, t_i) \geq u_i(\hat{t}_i, t_i)$ for all $\hat{t}_i \in T_i$.

As we focus on mechanisms that are BIC, we often do not distinguish between reported types ($\hat{t}_i$) and true types (t_i).

Definition 3.3. For a BIC mechanism, the *expected social utility* is $\sum_i \mathbf{E}_{t_i \sim G_i} [u_i(t_i, t_i)]$, the sum of the agents' expected utility in the truthful reporting equilibrium. An *optimal social utility mechanism*

is a non-audited BIC mechanism that maximizes the expected social utility. The term *optimal social utility auditing mechanism* refers to an audited BIC mechanism that satisfies the same criterion.

3.3 The Unit Demand Setting and Auction Theory

Optimal non-audited mechanisms are well-understood in the case that all agents have unit demand, i.e., d_i is a Bernoulli random variable, taking value 0 or 1 for all i . To see this, note that we may identify the type t_i of agent i as the probability with which d_i is 1; then agent i 's expected value for being allocated a unit of good is precisely t_i , and this allows t_i to act as the *value* in a classical single-parameter auction. Therefore, classical results from auction theory translate to statements on non-audited mechanisms in the unit-demand setting.

LEMMA 3.4 (MYERSON [22]). *A non-audited mechanism $\mathcal{M} = (\mathbf{x}, \mathbf{p})$ in the unit-demand setting is BIC if and only if the following two conditions hold:*

Monotonicity: *For all i , the interim allocation rule $x_i(\hat{t}_i)$ is monotonically increasing in $\hat{t}_i$.*

Payment identity: *For all i , the interim payment rule p_i is*

$$p_i(\hat{t}_i) = \hat{t}_i x_i(\hat{t}_i) - \int_0^{\hat{t}_i} x_i(t) dt + C \quad (\text{MPI})$$

for some constant C .

Given our restriction to non-negative payments and the objective of maximizing social utility, we always take the constant C in (MPI) to be 0.

Since agents' types are now defined over the reals, the type distribution G_i admits a CDF and a density function (PDF), denoted as G_i and g_i respectively.

LEMMA 3.5 (HARTLINE AND ROUGHGARDEN [19]). *Let $\mathcal{M} = (\mathbf{x}, \mathbf{p})$ be a BIC non-audited mechanism. Then the expected utility of agent i is*

$$\mathbf{E}_t [t_i \cdot x_i(\mathbf{t}) - p_i(\mathbf{t})] = \mathbf{E}_t [\varphi_i(t_i) \cdot x_i(\mathbf{t})], \quad (1)$$

where $\varphi_i(t_i)$ is $\frac{1-G_i(t_i)}{g_i(t_i)}$.

4 AUDITING MECHANISMS FOR UNIT-DEMAND AGENTS

In this section, we demonstrate the power of auditing in the unit-demand setting and characterize the social utility-optimal auditing mechanism. As discussed in Section 3, the types in this setting can be represented as $t_i = \Pr[d_i = 1]$. We can now write the interim utility for an allocation in the unit demand setting as

$$\begin{aligned} u_i(\hat{t}_i, t_i) &= \mathbf{E}_{d_i \sim t_i, t_{-i} \sim G_{-i}} [\min(d_i, x_i(\hat{t}_i, \mathbf{t}_{-i})) - p_i((\hat{t}_i, \mathbf{t}_{-i}), \min(d_i, x_i(\hat{t}_i, \mathbf{t}_{-i})))] \\ &= t_i \cdot x_i(\hat{t}_i) - t_i \cdot p_i(\hat{t}_i | 1) - (1 - t_i) \cdot p_i(\hat{t}_i | 0). \end{aligned}$$

Having to optimize over two payment terms makes characterizing the optimal social utility auditing mechanism difficult. The following theorem lets us focus on the simpler case in which agents never get charged when they have observed demand 1.

THEOREM 4.1. *For any collection T of type spaces and priors G over T , there exists an optimal social utility auditing mechanism such that $p_i((t_i, \mathbf{t}_{-i}), 1) = 0$ for all $\mathbf{t} \in T$ and all $i \in [N]$.*

PROOF. We show that any auditing mechanism $\mathcal{M} = (\mathbf{x}, \mathbf{p})$ with $p_i((t_i, \mathbf{t}_{-i}), 1) > 0$ for some set of types can instead be written as a mechanism $\mathcal{M}^* = (\mathbf{x}^*, \mathbf{p}^*)$ such that $p_i((t_i, \mathbf{t}_{-i}), 1) = 0$ for all $i \in [N]$ and all $\mathbf{t} \in T$ that obtains the same social utility.

We define new allocation and payment rules $\mathbf{p}^*$ and $\mathbf{x}^*$. For all i and for any $\mathbf{t}$ such that $p_i((t_i, \mathbf{t}_{-i}), 1) > 0$ we set $\mathbf{p}_i^*((t_i, \mathbf{t}_{-i}), 1) = 0$, and set the allocation to $\mathbf{x}_i^*(\mathbf{t}) = x_i(\mathbf{t}) - p_i((t_i, \mathbf{t}_{-i}), 1)$.

For any other type profile and agent, their allocations and payments remain the same. The resulting allocation is clearly feasible and non-negative. Finally, since $t_i(x_i^*(t'_i, t_{-i}) - p_i^*((t'_i, t_{-i}), 1)) = t_i(x_i(t'_i, t_{-i}) - p_i((t'_i, t_{-i}), 1))$ for every $t_i, t'_i \in T_i$ and any t_{-i} , there are no changes to the utility for any types and hence we maintain incentive compatibility. $\square$

4.1 Payment Reduction

So far we have shown that we may without loss of generality focus on mechanisms that only charge the agents that do not utilize an allocated item. We will refer to such auditing mechanisms as *waste-not-pay-not mechanisms*. As with traditional mechanism design, in waste-not-pay-not mechanisms, the payment rule is determined by the allocation rule, which allows us to fully characterize these mechanisms.

LEMMA 4.2. *Every waste-not-pay-not mechanism satisfies BIC constraints if and only if for each agent i , the following two conditions hold:*

- (1) *The interim allocation rule x_i is monotone non-decreasing.*
- (2) *The expected payment for reporting $\hat{t}_i$ when the observed demand is 0 is*

$$p_i(\hat{t}_i \mid 0) = C + \frac{\hat{t}_i \cdot x_i(\hat{t}_i)}{1 - \hat{t}_i} - \int_0^{\hat{t}_i} \frac{x(v)}{(1 - v)^2} dv, \quad (\text{API})$$

where $C \geq 0$ is a constant.

As our objective is to maximize social utility, without loss of generality we may take the C in (API) to be 0. The proof of Lemma 4.2 follows the same steps as the proof of Myerson's lemma [22], so we relegate it to the full paper.

We now derive some immediate consequences of Lemma 4.2. We will quantify in Theorem 4.6 the improvement in the social utility brought by auditing. To this end, we need to briefly review utility maximization without auditing.

Definition 4.3 (*M-unit α -lottery – Hartline and Roughgarden [19]*). The *M-unit α -lottery* allocates to agents with type at least α at price α . If there are more than M such agents, the winning agents are selected uniformly at random.

Definition 4.4 (*M-unit (α, β) -lottery – Hartline and Roughgarden [19]*). The *M-unit (α, β) -lottery* is the following mechanism: Let A and B denote the set of agents with reported types in the range (α, ∞) and $(\beta, \alpha]$, respectively.

- (1) If $|A| \geq M$, run an *M-unit α -lottery* on the $|A|$ agents with the highest types.
- (2) If $|A| + |B| \leq M$, sell to the $|A| + |B|$ agents with the highest types at price β .
- (3) Otherwise, run an $(M - |A|)$ -unit β -lottery for the agents in B , and allocate a unit to each agent in A at the price dictated by Myerson's payment identity (MPI).

THEOREM 4.5 (19). *Given any type distributions $G_1, \dots, G_n$, there is a mechanism that maximizes the expected social utility among all symmetric, BIC non-audited mechanisms, such that for every type profile $\mathbf{t}$, the mechanism's allocation rule is given by an M-unit (α, β) -lottery, for some α and β .*

For any (α, β) -lottery, we quantify the utility gained from only changing the associated payments with auditing (API) while keeping the same allocation rule. We discuss the utility gain in the three cases of Definition 4.4.

THEOREM 4.6. *For any (α, β) -lottery and truthfully reported types $\mathbf{t}$, a waste-not-pay-not mechanism with the same allocation rule but with payments determined by (API) generates a social utility that is weakly higher than that of the mechanism without auditing. If we let Δ be the difference in social utility between the two mechanisms, then:*

(1) if $|A| \geq M$,

$$\Delta = \sum_{i \in A} \frac{M}{|A|} \cdot \frac{\alpha(t_i - \alpha)}{(1 - \alpha)};$$

(2) if $|A| + |B| \leq M$,

$$\Delta = \sum_{i \in A \cup B} \frac{\beta(t_i - \beta)}{(1 - \beta)};$$

(3) if $|A| < M < |A| + |B|$,

$$\Delta = \sum_{i \in A} \frac{|B| + |A| - M}{|B|} \cdot \frac{\alpha(t_i - \alpha)}{(1 - \alpha)} + \sum_{i \in A \cup B} \frac{M - |A|}{|B|} \cdot \frac{\beta(t_i - \beta)}{(1 - \beta)}.$$

Moreover, whenever the non-audited (α, β) -lottery generates positive utility, the audited mechanism generates strictly higher utility.

Before proving the theorem, we provide some intuition for the reduction in the expected payment made possible by auditing. Consider agent i with type t_i in a second-price auction, facing α as the highest reported type from the other agents. In a waste-not-pay-not auction, a payment is only charged to agent i if i is allocated an item (i.e., $t_i \geq \alpha$) and do not use it. This occurs with probability $1 - t_i$; whenever this payment is less than $\frac{\alpha}{1-t_i}$, in expectation the agent will pay less than in the second-price auction without auditing. The payment should also be high enough so that lower types have no incentive to misreport and get allocated an item; therefore this payment should be at least $\frac{\alpha}{1-t'_i}$ for any t'_i that does not receive an item. The highest such t'_i is simply α , and therefore the amount the agent expects to pay in the audited auction is $(1 - t_i) \frac{\alpha}{1-\alpha}$.²

We now give a full proof for case 1 of Theorem 4.6, and relegate the rest to the full paper.

PROOF OF THEOREM 4.6, PART (1). Since the allocations remain unchanged in the audited mechanism, and since the audited mechanism is waste-not-pay-not, it suffices to calculate how much less is paid by an agent who is allocated an item. For an agent i with $t_i > \alpha$, let p_i be their payment in the non-auditing mechanism and let p_i^A be their payment in the auditing mechanism when they are allocated but the observed demand is 0 (recall if the observed demand is 1 there is no payment; also recall that the type profile $\mathbf{t}$ is fixed in the theorem statement). The item is given with equal probability to each agent in A , so by the payment identity $p_i = \frac{M}{|A|} \alpha$. Using our audited payment identity (API), we may calculate p_i^A :

$$\begin{aligned} p_i^A &= \frac{M \cdot t_i}{|A|(1 - t_i)} - \int_{\alpha}^{t_i} \frac{M}{|A|(1 - v)^2} dv \\ &= \frac{M \cdot t_i}{|A|(1 - t_i)} - \frac{M}{|A|(1 - t_i)} + \frac{M}{|A|(1 - \alpha)} \\ &= \frac{p_i M}{|A|(1 - \alpha)}. \end{aligned}$$

We now compute the difference in expected payment for each agent as $p_i - (1 - t_i)p_i^A = \frac{M}{|A|} \frac{\alpha(t_i - \alpha)}{(1 - \alpha)}$, and summing over the agents gives the expression in the theorem. $\square$

²This example of modified second-price can also be derived via the mechanism/approach in Ma et al. [21] when the center is only concerned with whether or not the item is consumed.

4.2 Optimal Auditing Mechanism

The previous results demonstrate that whenever the optimal social utility mechanism charges a payment, auditing lets us increase social utility without changing the allocation. This leaves the question as to whether mechanisms with auditing can improve the social utility further by using different allocation rules. Hartline and Roughgarden [19] showed that the optimal social utility mechanism is a *virtual utility maximizer* (see Section 3.3 for a review). In this section we derive a new formulation of virtual utility in the auditing setting and then show an example where this results in a different utility-maximizing allocation rule from its non-auditing counterpart.

THEOREM 4.7. *A BIC waste-not-pay-not mechanism with interim allocation rule $\mathbf{x}$ achieves expected social utility*

$$\sum_i \mathbb{E}_{t_i} \left[x_i(t_i) \cdot \frac{(1 - G_i(t_i))(1 - \mathbb{E}[b \mid b \geq t_i])}{g_i(t_i)(1 - t_i)^2} \right]. \quad (2)$$

This theorem allows us to associate with each type t_i a *virtual utility* $x_i(t_i) \cdot \frac{(1 - G_i(t_i))(1 - \mathbb{E}[b \mid b \geq t_i])}{g_i(t_i)(1 - t_i)^2}$, and maximizing social utility boils to designing monotone allocation rules that maximize (2).

The proof of Theorem 4.7 is similar to Myerson's derivation of virtual values for revenue maximization. We start by computing the expected value of $(1 - t_i)p(t_i \mid 0)$.

LEMMA 4.8. *In a BIC waste-not-pay-not mechanism, the (ex-ante) expected payment agent i makes is*

$$\mathbb{E}_{t_i} [(1 - t_i)p_i(t_i \mid 0)] = \mathbb{E}_{t_i} \left[x_i(t_i) \left(t_i - \frac{(1 - G_i(t_i))(1 - \mathbb{E}[b \mid b \geq t_i])}{g_i(1 - t_i)^2} \right) \right].$$

The proof of this lemma follows similar steps to the derivation of virtual values in revenue maximization literature. We relegate the proof to the full paper.

We can now use the audited payment identity (API) from Lemma 4.8 to prove Theorem 4.7.

PROOF OF THEOREM 4.7. The utility a mechanism generates can be written as

$$\sum_i \mathbb{E}_{t_i} [t_i \cdot x_i(t_i) - (1 - t_i)p_i(t_i \mid 0)]. \quad (3)$$

Using lemma 4.8, we can rewrite (3) as

$$\begin{aligned} & \sum_i \mathbb{E}_{t_i} \left[t_i \cdot x_i(t_i) - x_i(t_i) \left(t_i - \frac{(1 - G_i(t_i))(1 - \mathbb{E}[b \mid b \geq t_i])}{g_i(t_i)(1 - t_i)^2} \right) \right] \\ &= \sum_i \mathbb{E}_{t_i} \left[x_i(t_i) \cdot \frac{(1 - G_i(t_i))(1 - \mathbb{E}[b \mid b \geq t_i])}{g_i(t_i)(1 - t_i)^2} \right]. \quad \square \end{aligned}$$

We now show by an example that, even when the optimal non-auditing mechanism does not charge a payment, an auditing mechanism can still generate strictly higher social utility. In Theorem 4.9 we give a type distribution such that, when all agents' types are drawn i.i.d. from it, the optimal mechanism without auditing allocates the items uniformly at random without charging anyone (i.e., runs an M -unit 0-lottery), whereas with auditing the optimal mechanism is an M -unit α -lottery with $\mathbb{E}[\alpha] > 0$.

THEOREM 4.9. *When the center has one unit to allocate, there exists a type distribution G such that when all agents' types are drawn i.i.d. from G , the optimal social utility non-auditing mechanism allocates the items uniformly randomly without charging any agent; the optimal social utility auditing mechanism generates strictly more utility in expectation.*

PROOF. Let G be the exponential distribution with $\lambda = 1$ curtailed at 1 and renormalized. Namely, $g(t) = \frac{e^{-t}}{1-e^{-1}}$ and $G(t) = \frac{1-e^{-t}}{1-e^{-1}}$. Consider $G_1 = \dots = G_n = G$, that is, all agents' types are drawn i.i.d. from G . Hartline and Roughgarden [19] showed that, in the i.i.d. setting, when the type distribution's hazard rate, $H = \frac{g}{1-G}$, is monotone increasing, the optimal social utility (non-auditing) mechanism allocates the item uniformly at random with no charge. The hazard rate for G is $\frac{e^{-t_i}}{e^{-t_i}-e^{-1}}$, which is indeed a monotone increasing function; hence, the optimal social utility mechanism gives the item away at random for free.

We now consider the optimal social utility auditing mechanism. From Theorem 4.7 and our characterization of BIC waste-not-pay-not mechanisms (Lemma 4.2), we see that if $\frac{(1-G_i(t_i))(1-E[b | b \geq t_i])}{g_i(t_i)(1-t_i)^2}$ is monotone increasing in t_i then the optimal social utility auditing mechanism is the one that solves

$$\max_{\mathbf{x}} \sum_i \mathbb{E}_{t_i} \left[x_i(t_i) \cdot \frac{(1-G_i(t_i))(1-E[b | b \geq t_i])}{g_i(t_i)(1-t_i)^2} \right].$$

When we compute the virtual utilities for our distribution G , we find that these values are strictly increasing in t_i (the remainder of the proof along with the plot is included in the full paper). Therefore, the optimal social utility auditing mechanism sells the item to the agent with the highest type and charges the price dictated by Lemma 4.2. This allocation rule is a 1-unit α -lottery, where α is the selling price in a second-price auction. We also see that giving away the item for free is not a solution to this maximization problem. Therefore, the auditing mechanism must achieve strictly higher utility in this setting. $\square$

5 BEYOND UNIT DEMAND

When we go beyond unit-demand settings, the agents' preferences are multi-dimensional. Beyond the VCG auction, which maximizes social welfare, much less is known about mechanisms that optimize other objectives such as revenue or social utility, even without the complication of auditing. We show in this section that auditing still helps us increase social utility. We propose two variants of the VCG auction which keep the VCG allocation but demonstrably reduce the payments made by all agents. The first variant works without any distributional assumptions, and uses payments that are similar to the ones derived in the unit-demand setting. The second variant enhances utility further but requires an additional condition on the demand distributions.

5.1 The Audited VCG (AVCG) Mechanism

In this section we will describe a variant of the VCG mechanism that rescales payments in a way similar to the auditing payments in Section 4; we refer to this audited VCG mechanism as AVCG. First, recall that the VCG mechanism allocates the items to maximize social welfare. We show that the VCG allocation is given by a greedy procedure: each agent i with type t_i , given that they have been assigned k units of items, has marginal value $1 - F_{t_i}(k)$ for being assigned an additional item, and this decreases as k grows; therefore, given the agents' types, the problem boils down a multi-unit auction where agents have decreasing marginal values. It is straightforward to see that, to maximize welfare, one should allocate the M items in a greedy fashion: allocating each item sequentially to an agent whose current marginal value is highest, where the marginal value for agent i already allocated k items is $1 - F_{t_i}(k)$. Let $\mathbf{x}^{\text{VCG}}$ denote the resulting allocation rule. The VCG mechanism then charges each agent i the externality i causes. In this setting, the payment $p_i^{\text{VCG}}(\mathbf{t})$ is thus the sum of the $(M - x_i^{\text{VCG}}(\mathbf{t}) + 1)$ -st through M -th highest marginal utilities that do not belong to agent i .

One way to formalize the preceding intuition is as follows: if we let $\mu_i^k(t_i)$ be $1 - F_{t_i}(k)$, agent i 's marginal value for the $(k + 1)$ -st unit, and let $v_i^k(\mathbf{t}_{-i})$ be the $(M - k)$ -th highest marginal utility

not belonging to agent i , then agent i receives the $(k + 1)$ -st unit if and only if $\mu_i^k \geq v_i^k$ (up to tie-breaking). Thus the payment of agent i in the VCG mechanism can be written in terms of v as $p_i^{\text{VCG}}(\mathbf{t}) = \sum_{k=0}^{x_i^{\text{VCG}}(\mathbf{t})-1} v_i^k(\mathbf{t}_{-i})$.

We now describe our variants of the VCG mechanism using auditing. Both variants keep the VCG allocation rule, while decreasing the expected payment made by each agent.

Definition 5.1. The *audited VCG* (AVCG) mechanism uses the allocation rule $\mathbf{x}^{\text{VCG}}$ of the VCG mechanism and has payment rule

$$p_i^{\text{AVCG}}(\mathbf{t}, d_i^{\text{obs}}) := \sum_{k=0}^{x_i^{\text{VCG}}(\mathbf{t})-1} \frac{\mathbf{1}_{d_i^{\text{obs}} \leq k}}{1 - v_i^k(\mathbf{t}_{-i})} v_i^k(\mathbf{t}_{-i}),$$

where $\mathbf{1}_{\mathcal{A}}$ denotes the indicator function for the event $\mathcal{A}$.

In both the VCG and the AVCG mechanisms, the payments can be thought of on a per-unit basis. In VCG, the payment for the k -th unit is $v_i^k(\mathbf{t}_{-i})$, regardless of realized demand. In AVCG, the agent pays nothing for the k -th unit if they use at least k units. However, if the agent is allocated a k -th unit, but does not use it, the agent pays $v_i^k(\mathbf{t}_{-i})/(1 - v_i^k(\mathbf{t}_{-i}))$; this value can be thought of as the VCG payment diluted by a factor of $1 - v_i^k(\mathbf{t}_{-i})$. Although the resulting payment in the case that the agent does not use all allocated items is larger than the VCG payment, we will show that the overall *expected* payment for any truthful agent is *at most* that of VCG. We show moreover that the expected payment for any non-truthful agent is always *at least* that of VCG; the intuition here is that agents who misreport and claim higher demand than what they expect are punished more heavily for failing to use allocated units.

Theorems 5.2 and 5.3 formalize the preceding claims. A key step of the proof involves analyzing the expected payment for agent i ; we can write this expected payment for agent i with type t_i and reporting $\hat{t}_i$, while the other agents report a profile of $\mathbf{t}_{-i}$, as

$$\mathbb{E}_{d_i \sim t_i} [p_i^{\text{AVCG}}((\hat{t}_i, \mathbf{t}_{-i}), d_i^{\text{obs}})] = \sum_{k=0}^{x_i^{\text{VCG}}(\hat{t}_i, \mathbf{t}_{-i})-1} \frac{1 - \mu_i^k(t_i)}{1 - v_i^k(\mathbf{t}_{-i})} v_i^k(\mathbf{t}_{-i}).$$

THEOREM 5.2. *The AVCG mechanism is dominant strategy incentive compatible.*

PROOF. If, by misreporting their type t_i as $\hat{t}_i$, agent i changes the outcome and receives more items than they would be allocated under truthful reporting, then it must be true that $\mu_i^k(t_i) \leq v_i^k(\hat{\mathbf{t}}_{-i})$ for each additional unit k they receive. It follows that $\frac{1 - \mu_i^k(t_i)}{1 - v_i^k(\hat{\mathbf{t}}_{-i})} \geq 1$, in which case the expected payment for the k -th unit is $\frac{1 - \mu_i^k(t_i)}{1 - v_i^k(\hat{\mathbf{t}}_{-i})} v_i^k(\hat{\mathbf{t}}_{-i}) \geq \mu_i^k(t_i)$. This shows that the agent incurs non-positive expected returns for all units k for which they overreported and were allocated. Hence the agent is weakly better off reporting truthfully.

On the other hand, if by misreporting, agent i changes the outcome and receives fewer items than they would be allocated with truthful reporting, we claim they would also be weakly worse off. To see this, note that agent i loses utility in expectation for each unit lost, since the expected utility from any allocated unit under truthful reporting is always non-negative. Furthermore, their expected payments for the units that they are allocated remain unchanged from the truthful setting, since payments for agent i are determined entirely by the bids of the other agents. Thus agent i will never benefit from changing the outcome to receive fewer items.

Agent i does not benefit from changing the outcome away from that of truthful play, regardless of what other agents do, so the AVCG mechanism is dominant strategy incentive compatible. $\square$

THEOREM 5.3. *In comparison to VCG, the AVCG mechanism satisfies the following:*

- (1) *For each agent i with type t_i , for any reported type profile $\mathbf{t}_{-i}$ of the other agents, the expected payment made by agent i with truthful bidding is at most the agent's VCG payment.*
- (2) *For any type profile $\mathbf{t}$ with which agent i obtains strictly positive utility in the VCG mechanism, the agent obtains strictly higher expected utility in the AVCG mechanism.*

PROOF. Notice that the condition $\mu_i^k(t_i) \geq v_i^k(\mathbf{t}_{-i})$ for allocating a k -th unit to agent i in the VCG allocation rule implies $\frac{1-\mu_i^k(t_i)}{1-v_i^k(\mathbf{t}_{-i})} v_i^k(\mathbf{t}_{-i}) \leq v_i^k(\mathbf{t}_{-i})$. Summing over k , we see that the expected AVCG payment is always at most the VCG payment. Furthermore, if for any i and k , $\mu_i^k(t_i) > v_i^k(\mathbf{t}_{-i})$, then the inequality is strict; this is exactly the condition under which VCG extracts positive utility. Thus, if VCG extracts positive utility, the AVCG mechanism yields strictly higher utility. $\square$

Our construction of the AVCG mechanism shows that under very general assumptions, auditing lets us outperform the VCG mechanism in terms of utility maximization. The AVCG mechanism accomplishes this goal while preserving key properties of VCG, including the surplus-maximizing allocation rule and dominant strategy incentive compatibility, without making any assumptions about demand distributions.

5.2 Stochastic Dominance

We now describe our second variant of the VCG mechanism, which further improves utilities with auditing, but requires an assumption on types—specifically, that the types in each agent's type space admit an ordering by stochastic dominance.

Definition 5.4. We say an agent i 's types admit a *first-order stochastic dominance ordering (FSDO)* if there exists an ordering on T_i such that if $t_i > t'_i$ then $\mu_i^k(t_i) \geq \mu_i^k(t'_i)$ for any demand k .

We must also define a mapping that allows us to relate the marginal values of opponents with the marginal values of each agent's other potential types. We define a mapping for all marginal values v , and $j, k \in \mathbb{N}_{\geq 0}$, $\psi_i(v, k, j) = \sup_{\{t_i: \mu_i^j(t_i) \leq v\}} \mu_i^k(t_i)$. Note for all values of k the type t_i that satisfies this condition does not change, only the outputted marginal value. Intuitively, this type t_i can be thought of as the highest type who would not be allocated j items in the presence of the unallocated opponent's marginal v .

Definition 5.5. For agents whose type spaces each admit a first-order stochastic dominance ordering, the *stochastic-dominance audited VCG (SD-AVCG)* mechanism has allocation rule $\mathbf{x}^{\text{VCG}}$ the same as the VCG mechanism, and payment rule

$$p_i((t_i, \mathbf{t}_{-i}), d_i^{\text{obs}}) = \sum_{j=0}^{x_i^{\text{VCG}}(\mathbf{t})-1} \psi_i(v_i^j(\mathbf{t}_{-i}), j, j) \min_{k \leq x_i^{\text{VCG}}(\mathbf{t})} \left[\frac{\mathbf{1}_{d_i^{\text{obs}} \leq k}}{1 - \psi_i(v_i^j(\mathbf{t}_{-i}), k, j)} \right].$$

The key difference between SD-AVCG and AVCG is that by framing prices in terms of an agent's own marginal values, leveraging FSDO, we are able to construct a truthful payment rule that conditions over any of the possible values of d_i^{obs} . We will prove a result analogous to Theorems 5.2 and 5.3. Again we make use of the expected payment,

$$\mathbb{E}_{d_i \sim t_i} [p_i((t_i, \mathbf{t}_{-i}), d_i^{\text{obs}})] = \sum_{j=0}^{x_i^{\text{VCG}}(\mathbf{t})-1} \psi_i(v_i^j(\mathbf{t}_{-i}), j, j) \min_{k \leq x_i^{\text{VCG}}(\mathbf{t})} \left[\frac{1 - \mu_i^k(t_i)}{1 - \psi_i(v_i^j(\mathbf{t}_{-i}), k, j)} \right].$$

THEOREM 5.6. *If all agents' types admit first-order stochastic dominance orderings, then the SD-AVCG mechanism is dominant-strategy incentive compatible. Furthermore:*

- (1) *For each agent i with type t_i , for any reported type profile $\mathbf{t}_{-i}$ of the other agents, the expected payment made by agent i with truthful bidding is at most the agent's AVCG payment.*
- (2) *For any type profile $\mathbf{t}$ with which agent i obtains strictly positive utility in the VCG mechanism, the agent obtains strictly higher expected utility in the SD-AVCG mechanism.*

PROOF. For dominant-strategy incentive compatibility, the proof follows the mostly same steps as the proof of Theorem 5.2. The part that differs from AVCG is the minimization function. Here we use the fact FSDO guarantees for any $j \leq x^{\text{VCG}}(t_i)$ and for all k , $\frac{1-\mu_i^k(t_i)}{1-\psi_i(v_i^j(\mathbf{t}_{-i}), k, j)} \leq 1$, while the converse is true for $j > x^{\text{VCG}}(t_i)$.

For statement (1), recognize that for any type profile $\mathbf{t}_{-i}$, the mapping $\psi_i(v_i^j(\mathbf{t}_{-i}), j, j)$ is bounded above by $v_i^j(\mathbf{t}_{-i})$. Therefore the first term in the payment of SD-AVCG is trivially smaller than in AVCG. For the second term, the “scaling factor” used in AVCG is always a potential solution to the minimization in SD-AVCG. Therefore, the SD-AVCG payment is at most the AVCG payment.

Statement (2) follows from combining Theorem 5.3 with statement (1). $\square$

A simple example of when SD-AVCG extracts more utility than AVCG is when every type $t'_i < t_i$ has a demand value that is smaller than the smallest value in the support of t_i (i.e. as types increase in the ordering so does the lower bound of their support). In this case, SD-AVCG is guaranteed to charge an expected payment of 0 for all items by conditioning on the 0-probability event, but this is not always possible with AVCG.

6 DEBT MECHANISMS FOR REPEATED ALLOCATION

An important distinction between auditing mechanisms and more standard mechanisms is that under auditing, agents' payments occur after the allocation has been received (in order to allow time for the audit to take place). Thus, it is perhaps especially natural to consider the effect auditing can have in repeated settings in which agents interact with the mechanism more than once.

In this section, we describe *debt mechanisms*, a novel non-monetary mechanism design for settings with repeated interactions. Debt mechanisms give a way to extend any static mechanism into a repeated mechanism that does not require monetary transfers at all. In addition, for truthful single-round mechanisms, the resulting debt mechanism achieves social welfare equal to the social utility of the single-round mechanism—thus making auditing a powerful tool.

Our reduction relies on a variant of the classic observation that repetition allows us to use future allocative penalties to dis-incentivize misreporting (at least for sufficiently patient players). Agents found to have misreported become “indebted” over time; they are eventually docked allocations until their debt is cleared. Using future periods this way makes auditing mechanisms particularly practical: they can be used in settings (like food allocation) in which centers (food banks) and agents (food pantries) interact frequently, but having any sort of monetary transfer might be inappropriate.

6.1 Setup

A *debt mechanism* $\mathcal{M}_d$ is a repeated mechanism that runs for infinitely many rounds; it can be uniquely described by a single-round quasilinear mechanism $\mathcal{M} = (\mathbf{x}, \mathbf{p})$, an allocation length l , and a debt rate for each agent $\mathbf{r} \in \mathbb{R}^N$, where r_i is the debt rate for agent i . For debt mechanisms, we again make the assumption of non-negative payments for $\mathcal{M}$, i.e. $p_i \geq 0$ always.³

³Note that with this construction, we implicitly assume that auditing occurs every round, since it is a component of our single-round mechanisms. In the future, it would be interesting to understand the extent to which our approach and results extend the settings with less frequent (or stochastic) auditing.

A debt mechanism records each agent's total debt to the mechanism over time, and at any given round i , each agent can be in one of two states:

- (1) **Allocation:** In an *allocation round*, the agent participates in the single-round mechanism $\mathcal{M}$ and is allocated according to the allocation rule defined by $\mathcal{M}$; any payment incurred is added to the agent's debt.
- (2) **Punishment:** In a *punishment round*, the agent is allocated nothing and is charged no payment; the agent's debt is reduced according to the debt update rule described below.

We call an allocation round that occurs directly after a punishment round an *allocation start round*. The *allocation length* l is the number of guaranteed consecutive allocation rounds starting from an allocation start round. For example, if $l = 3$ and round 2 is an allocation start round, then rounds 3 and 4 are guaranteed to be allocation rounds. The *debt rate* r_i of agent i is the amount of debt that is subtracted from agent i 's total debt before every round whose probability of being a punishment round is non-zero, resetting to 0 if negative. The type of round an agent is in at any point in time is determined by the amount of debt they owe to the mechanism.

We now carefully describe the relationship between debt and the probability of being in an allocation or punishment round. We define an indicator variable a_i^j for each agent i and each round j such that $a_i^j = 1$ if round j is an allocation round for agent i and $a_i^j = 0$ if round j is a punishment round for agent i . We can write the probability that $a_i^j = 0$ in terms of the debt left to be paid off by that agent. Let s be the most recent allocation start round such that $s + l \leq j$. Then

$$\Pr \left[a_i^j = 0 \right] = \min \left(\max \left(\frac{\sum_{k=s}^{s+l} p_i(\sigma^k) - (j - s - l)r_i}{r_i}, 0 \right), 1 \right), \quad (4)$$

where σ^k is the type the agent reported at round k .

Intuitively, (4) says that the probability an agent is in a punishment round is 1 if their debt is larger than the debt rate and it has been at least l rounds since their last allocation start round. Furthermore, they are in a punishment round with probability equal to their debt divided by the debt rate if their debt is less than the debt rate and it has been at least l rounds since the last allocation start round. In both of these cases, the agent's debt is decreased by the minimum of their debt rate and their remaining debt at the start of the round. Otherwise an allocation round is guaranteed—either by the guarantee given by allocation length or by having 0 debt.

6.2 Stationary Strategies

In this section we describe the strategy space available to the agents and characterize a smaller set of optimal strategies. We begin by considering the case with a single agent. For ease of notation, we drop the subscript referring to the agent number. At round j , an agent has type t^j drawn from their prior distribution G . If $a^j = 1$, an agent must report a type to the mechanism $\mathcal{M}$ according to some strategy. This strategy may depend not only on the agent's type at round j , but also on the amount of time since the mechanism began and the history of all allocations and payments made until this round. We define the history up to round j as $H^j := \{(a^1 x(t^1), \dots, a^{j-1} x(t^{j-1})), (a^1 p(t^1), \dots, a^{j-1} p(t^{j-1}))\} \cup \{a^j\}$. Let $\mathcal{H}$ be the space of all potential histories. A strategy for an agent is simply a map $\sigma : \mathcal{H} \times \mathbb{Z}_{>0} \times T \rightarrow T$, mapping tuples of (history, round number, type) to a reported type.

We assume agents choose a strategy that gives them the most (expected) value, where value is a (potentially randomized) function of the agent's type and the allocation received that is additive over time. As agents must reason about an infinite time horizon, we assume each agent chooses their strategy based on the following criterion.

Definition 6.1 (Overtaking Criterion). Let $V^j(\sigma^j)$ be the amount of value an agent receives for playing strategy σ at round j . A strategy σ is *weakly preferred* to a strategy $\hat{\sigma}$ if and only if $\liminf_{n \rightarrow \infty} (\sum_{j=1}^n V^j(\sigma^j) - \sum_{j=1}^n V^j(\hat{\sigma}^j)) \geq 0$.

For a given debt mechanism $\mathcal{M}_D = ((x, p), l, r)$, we can rewrite $V^j(\sigma^j) = V(t^j, x(\sigma^j))$. Therefore, an optimal strategy σ^* is a strategy such that for all σ' ,

$$\liminf_{n \rightarrow \infty} \left(\sum_{j=1}^n \mathbb{E} [a^j \cdot V(t^j, x(\sigma^*(H, j, t^j)))] - \sum_{j=1}^n \mathbb{E} [a^j \cdot V(t^j, x(\sigma'(H, j, t^j)))] \right) \geq 0.$$

We now introduce the concept of stationary strategies. Informally, a strategy is said to be *stationary* if it does not depend on time or history. The formal definition is as follows.

Definition 6.2. A strategy σ is *stationary* if and only if, for all types $t \in T$, histories $H, H' \in \mathcal{H}$, and rounds $k, k' \in \mathbb{Z}_{>0}$, we have $\sigma(H, k, t) = \sigma(H', k', t)$.

For ease of notation we write stationary strategies as maps $\sigma_s : T \rightarrow T$. We now show that for any debt mechanism there exists an optimal strategy σ^* that is stationary. We first show this result for the case of allocation length $l = 1$.

THEOREM 6.3. *For any debt mechanism $\mathcal{M}_D$ with allocation length $l = 1$, there exists a stationary strategy σ^S such that for all other strategies $\hat{\sigma}$,*

$$\liminf_{n \rightarrow \infty} \left(\sum_{j=1}^n \mathbb{E} [a^j \cdot V(t^j, x(\sigma_s(t^j)))] - \sum_{j=1}^n \mathbb{E} [a^j \cdot V(t^j, x(\hat{\sigma}(H, j, t^j)))] \right) \geq 0.$$

PROOF SKETCH. We begin by showing there exists an optimal strategy whose mapping at round k is independent of its history. We then use history-independence to show that any sequence of strategies at time k' can be replaced by the sequence of strategies starting at time k without loss of utility. The remaining details are included in full paper. $\square$

A nearly identical argument to that for Theorem 6.3 shows that there exists an optimal stationary strategy regardless of the allocation length.⁴ Formally, we have:

THEOREM 6.4. *For any debt mechanism $\mathcal{M}_D$, there exists a stationary strategy σ_S such that for all other strategies σ' , $\liminf_{n \rightarrow \infty} (\sum_{j=1}^n \mathbb{E}[a^j \cdot V(t^j, x(\sigma_S(t^j)))] - \sum_{j=1}^n \mathbb{E}[a^j \cdot V(t^j, x(\sigma'(H, j, t^j)))] \geq 0$*

PROOF SKETCH. The proof follows the same steps as the proof of Theorem 6.3 with two small differences. The first difference is that we only prove history-independence for any two histories that are the same depth into an interval of allocation rounds (e.g., any histories which are the second consecutive allocation round). Therefore, when we prove time-independence, instead of just conditioning over whether or not round k is an allocation round, we must also condition over the l possible positions k could be within an allocation round. $\square$

The preceding results show that for any debt mechanism, agents can arrive at optimal strategies by only considering strategies that rely solely on their present-period types.

⁴We omit the full proof for brevity.

6.3 Truthfulness is Optimal

We now introduce a special case of debt mechanisms we call *truthful debt mechanisms*. We say that a debt mechanism $M_D = (\mathcal{M}, r, l)$ is *truthful* if its single-round mechanism $\mathcal{M} = (x, p)$ satisfies BIC constraints and $r = \mathbb{E}[V(t, x(t)) - p(t)]$ for each agent. We show that for any truthful debt mechanism, an optimal stationary strategy is one that maps every type to itself, i.e., is a truthful strategy. By the existence of optimal strategies that are stationary, this also implies that the optimal strategy for such a mechanism is truthful.

By Theorem 6.3, we only need to compare being truthful with other stationary strategies. It thus suffices to compare the average value each strategy achieves, as no stationary strategies can differ by only a finite amount in the limit.⁵ To compute the expected average value obtained by a strategy, we need to compute both the expected proportion of rounds spent being punished and the expected value obtained during each allocation round. We can then compute the expected average value for a strategy as the product of the expected value from an allocation round and the expected proportion of time spent in allocation rounds.

We begin by computing the expected proportion of time spent in allocation rounds in the limit.

LEMMA 6.5. *The fraction of rounds that are allocation rounds when playing a stationary strategy σ converges in probability to $\frac{r}{r + \mathbb{E}_{t^j}[p(\sigma(t^j))]}$.*

PROOF. For a single set of k allocation rounds, the expected number of rounds spent being punished is $\mathbb{E}[\sum_{j=1}^l \frac{p(\sigma(t^j))}{r}]$. We refer to a single continuous block of l allocation rounds as an *allocation interval*. We refer to an allocation interval along with the series punishment rounds it takes to return to zero debt as an *interval*. The number of intervals contained in the first n rounds can be described by the sequence of random variables $\{I_n\}_{n \geq 0}$. As there is some finite and constant maximum on the length of any given interval, we know that as n approaches ∞ , I_n approaches ∞ with probability 1.

We now compute the average number of punishment rounds in the first n rounds as n approaches ∞ . We first define indices $(\pi_1, \dots, \pi_{I_n})$, where π_i is the round at which the punishment for the i -th allocation interval ends. Since each debt term is drawn independently, the choice of these indices does not affect the expected value. Note that at the end of n rounds for any given n , the last punishment round may or may not be complete, so we represent its length as the random variable l_p . We can now define the average number of punishment rounds as a random variable Z and compute its expected value:

$$\mathbb{E}[Z] = \lim_{n \rightarrow \infty} \mathbb{E} \left[\frac{\sum_{j=1}^l \frac{p(\sigma(t^j))}{r} + \sum_{j=\pi_1}^{\pi_1+l-1} \frac{p(\sigma(t^j))}{r} + \dots + \sum_{j=\pi_{I_n-1}}^{\pi_{I_n-1}+l-1} \frac{p(\sigma(t^j))}{r}}{I_n} + \frac{l_p}{I_n} \right]. \quad (5)$$

As I_n approaches ∞ with probability 1, we may simplify (5) to

$$\mathbb{E}[Z] = \lim_{n \rightarrow \infty} \mathbb{E} \left[\frac{\sum_{j=1}^l \frac{p(\sigma(t^j))}{r} + \sum_{j=\pi_1}^{\pi_1+l-1} \frac{p(\sigma(t^j))}{r} + \dots + \sum_{j=\pi_{I_n-1}}^{\pi_{I_n-1}+l-1} \frac{p(\sigma(t^j))}{r}}{I_n} \right]. \quad (6)$$

By law of large numbers, we have $\lim_{n \rightarrow \infty} \mathbb{E}[Z] = \mathbb{E}[\sum_{j=1}^l \frac{p(\sigma(t^j))}{r}]$, with the convergence taking place in probability. Clearly, for n rounds the average number of allocation rounds is equal to l by

⁵If two stationary strategies' averages converge to different values, then their total utilities must grow arbitrarily far apart as $n \rightarrow \infty$.

the same logic as above. We can now compute the fraction of rounds that are allocation rounds as

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E} \left[\frac{l}{l + Z} \right] &= \frac{l}{l + \lim_{n \rightarrow \infty} \mathbb{E}[Z]} \\ &= \frac{l}{l + l \cdot \frac{\mathbb{E}_{t^j}[p(\sigma(t^j))]}{r}} \\ &= \frac{r}{r + \mathbb{E}_{t^j}[p(\sigma(t^j))]} \end{aligned}$$

Note that the first line of the above follows from the Continuous Mapping Theorem [25, Chapter 2], as Z converges in probability to a constant. $\square$

We now use the expression for the proportion of allocation rounds from Lemma 6.5 along with the expected value for an allocation round (6) to prove our main theorem.

THEOREM 6.6. *For any truthful debt mechanism, reporting truthfully is optimal. That is, for any strategy σ' , $\liminf_{n \rightarrow \infty} \sum_{j=1}^n \mathbb{E}[a^j \cdot V(t^j, x(t^j))] - \sum_{j=1}^n \mathbb{E}[a^j \cdot V(t^j, x(\sigma'(H, j, t^j)))] \geq 0$.*

PROOF. Let $Y := \mathbb{E}[V(t^j, x(\sigma(t^j)))]$ be the expected value of playing a stationary strategy σ in any allocation round. The expected value of the strategy in round j is then $Y \cdot a^j$, since a^j and the value from the j -th round are independent random variables. (Note that a^j is determined entirely by the outcomes of previous rounds, whereas the utility of the agent in round j is solely a function of the agent's type and the outcome of round j .) Then the average expected utility after n rounds, by Lemma 6.5, converges in probability to

$$Y \cdot \frac{1}{n} \sum_{j=1}^n a^j \rightarrow Y \cdot \frac{r}{r + \mathbb{E}_{t^j}[p(\sigma(t^j))]} = \frac{\mathbb{E}[V(t^j, x(\sigma(t^j)))]}{r + \mathbb{E}_{t^j}[p(\sigma(t^j))]} r.$$

Since $\mathcal{M}$ is BIC we know that the ratio

$$\frac{\mathbb{E}[V(t^j, x(\sigma(t^j)))]}{r + \mathbb{E}_{t^j}[p(\sigma(t^j))]} = \frac{\mathbb{E}[V(t^j, x(\sigma(t^j)))]}{\mathbb{E}[V(t^j, x(t^j)) - p(t^j)] + \mathbb{E}[p(\sigma(t^j))]} \leq 1$$

for any non-truthful strategy σ . Furthermore, this ratio is exactly 1 when playing truthfully. Thus, the identity mapping provides at least as much value as any other stationary strategy. $\square$

We can also now see that the truthful debt mechanism achieves expected utility equal to the expected utility of the single-round mechanism.

COROLLARY 6.7. *For any truthful debt mechanism, reporting truthfully gives social welfare equal to the social utility of the single-round mechanism.*

PROOF. Reporting truthfully gives an agent an expected value of r where r is defined as the expected utility the agent would receive in the single-round mechanism. $\square$

Notice that the last condition required for truthfulness in the repeated game is not in the same form as the BIC constraints. Instead, it only requires that the strategy is truthful ex ante in the agent's own type. In other words, the condition can be written as

$$\mathbb{E}_t[V(t, x(\sigma(t))) - p(\sigma(t))] \leq \mathbb{E}[V(t, x(t)) - p(t)] \quad (7)$$

for all strategies σ . However, (7) is equivalent to satisfying the BIC constraints.

LEMMA 6.8. *A mechanism $\mathcal{M}$ satisfies BIC constraints if (7) holds for all mappings σ .*

The proof of Lemma 6.8 is a simple application of IC constraints and is shown in the full paper.

6.4 Multi-Agent Debt Mechanisms

Our discussion so far in this section addressed mechanisms for incentivizing individual agents to reveal their types truthfully. In constructing a debt mechanism that works for arbitrarily many agents, one additional intricacy arises: ensuring that the type of round each agent's opponents are in (allocation or punishment) does not affect that agent's incentives. Luckily, round/agent incentive mismatch can be avoided with the following simple procedure: When agents are in punishment rounds, their type distributions are randomly sampled and given to the mechanism $\mathcal{M}$ as inputs. Any goods that would have been allocated to the agent are instead left unallocated; this makes agents ambivalent as to whether their opponents are in punishment or allocation rounds (because agents' priors over possible opponent types remains unchanged).

7 CONCLUSIONS AND FUTURE WORK

We have explored how auditing can help us maximize social welfare, as measured by total utility—and we found that auditing enables mechanisms that outperform the existing state of the art in several senses. In particular, auditing allows us to reduce the payments imposed by the optimal (non-auditing) mechanism in the unit-demand setting, and also leads to a new optimal mechanism for that setting. We showed two different strategies for reducing VCG payments in more general settings: one that makes no distributional assumptions, and a second that achieves even better outcomes when agents' potential types can be ordered by stochastic dominance. Finally, we showed how to use auditing to lift any static mechanism to a repeated allocation setting in which payments are not allowed, via a general construction for producing what we call debt mechanisms.

We hope that auditing will lead to ways of improving social good in other mechanism design contexts. Additionally, two variants of our framework seem particularly worthy of mention for potential consideration.

First, we note that our debt mechanisms must sometimes waste units of the good in order to punish agents. This may be unpalatable even in our lead applications—we could naturally think that food banks, for example, would be unwilling to allow in-demand food to spoil just for the sake of incentives. We might thus consider mechanisms that designate one or more “sink agents” who are only ever allocated goods in punishment rounds. It is clear that mechanisms with sink agents cannot always be optimal, because in some settings, the allocation to sink agents will be too small in equilibrium. However, it is also clear that settings exist in which introducing the concept of sink agents can help: e.g., if there is an agent whose prior distribution is weak enough that it is never allocated, social utility is increased by designating that agent as a sink.

Second, it may be worth reconsidering our restriction that auditing mechanisms may not impose negative payments—that is, we might want to allow mechanisms that sometimes pay agents. The restriction to negative payments was essential for our transformation of auditing mechanisms into debt mechanisms. Indeed, if net negative payments correspond to wasting units of the good, then *positive* payments would then have to correspond to *creating* additional units of the good. Nevertheless, relaxing the no-negative-payments condition would allow more powerful auditing mechanisms in settings that do allow monetary payments. Of course, utility can always be (trivially) increased without bound by increasing payments to agents without bound, so it is necessary to impose a budget balance constraint. If we allow *ex ante* budget balance, it is straightforward to achieve perfect utility maximization with 0 expected payments: run VCG, and then unconditionally pay agents a $1/N$ fraction of the mechanism's revenue under VCG. In future work, it is worth considering what can be achieved under an *ex post* budget balance constraint or by a detail-free mechanism (i.e., a mechanism that does not depend on the priors).

More broadly, it might be worth examining the extent to which other auditing formats—especially costly auditing; see Ben-Porath et al. [2]—have value for utility maximization, and if so, whether any of our insights carry over. And lastly, it may be worthwhile to think about auditing as part of the objective: organizations might, for example, prefer to minimize the number of audits necessary, all else equal.

REFERENCES

- [1] Santiago Balseiro, Huseyin Gurkan, and Peng Sun. 2017. Multi-agent Mechanism Design without Money. (2017). Working Paper.
- [2] Elchanan Ben-Porath, Eddie Dekel, and Barton L. Lipman. 2014. Optimal allocation with costly verification. *American Economic Review* 104, 12 (2014), 3779–3813.
- [3] Jeremiah Blocki, Nicolas Christin, Anupam Datta, Ariel D. Procaccia, and Arunesh Sinha. 2013. Audit Games. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*. 41–47.
- [4] Ruggiero Cavallo. 2006. Optimal decision-making with minimal waste: Strategyproof redistribution of VCG payments. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*. ACM, 882–889.
- [5] Richard E. Caves. 2003. Contracts between art and commerce. *Journal of Economic Perspectives* 17, 2 (2003), 73–83.
- [6] Sofia Ceppi, Ian Kash, and Rafael Frongillo. 2018. Partial Verification as a Substitute for Money. *arXiv:1812.07312* (2018).
- [7] Surajeet Chakravarty and Todd Kaplan. 2013. Optimal allocation without transfer payments. *Games and Economic Behavior* 77, 1 (2013), 1–20.
- [8] Geoffroy de Clippel, Victor Naroditskiy, Maria Polukarov, Amy Greenwald, and Nicholas R. Jennings. 2014. Destroy to save. *Games and Economic Behavior* 86 (2014), 392–404.
- [9] Piotr Dworczak and © Scott Duke Kominers © Mohammad Akbarpour. 2019. Redistribution through markets. (2019). Becker Friedman Institute Working Paper.
- [10] Albin Erlanson and Andreas Kleiner. 2015. Costly Verification in Collective Decisions. (2015). Working Paper.
- [11] Albin Erlanson and Andreas Kleiner. 2018. A note on optimal allocation with costly verification. (2018). Working Paper.
- [12] Thomas Giebe and Elmar Wolfstetter. 2008. License auctions with royalty contracts for (winners and) losers. *Games and Economic Behavior* 63, 1 (2008), 91–106.
- [13] Artur Gorokh, Siddhartha Banerjee, and Krishnamurthy Iyer. 2017. From Monetary to Non-Monetary Mechanism Design via Artificial Currencies. In *Proceedings of the 2017 ACM Conference on Economics and Computation, EC '17, Cambridge, MA, USA, June 26-30, 2017*. 563–564.
- [14] Mingyu Guo and Vincent Conitzer. 2009. Worst-case optimal redistribution of VCG payments in multi-unit auctions. *Games and Economic Behavior* 67, 1 (2009), 69–98.
- [15] Mingyu Guo and Vincent Conitzer. 2010. Optimal-in-expectation redistribution mechanisms. *Artificial Intelligence* 174, 5-6 (2010), 363–381.
- [16] Mingyu Guo, Vincent Conitzer, and Daniel M. Reeves. 2009. Competitive Repeated Allocation without Payments. In *Proceedings of the 5th International Workshop on Internet and Network Economics, Rome, Italy, December 14-18, 2009*. 244–255.
- [17] Philip Haile, Kenneth Hendricks, and Robert Porter. 2010. Recent US offshore oil and gas lease bidding: A progress report. *International Journal of Industrial Organization* 28, 4 (2010), 390–396.
- [18] Robert G. Hansen. 1985. Auctions with contingent payments. *American Economic Review* 75, 4 (1985), 862–865.
- [19] Jason D. Hartline and Tim Roughgarden. 2008. Optimal mechanism design and money burning. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*. 75–84.
- [20] Scott Duke Kominers and Alan Lam. 2018–2019. Feeding America (A) and (B). (2018–2019). Harvard Business School Case 818-130, Supplement 818-131, and Teaching Note 918-082.
- [21] Hongyao Ma, Reshef Meir, David C. Parkes, and James Zou. 2016. Contingent Payment Mechanisms to Maximize Resource Utilization. *arXiv:1607.06511* (2016).
- [22] Roger Myerson. 1981. Optimal Auction Design. *Mathematics of Operations Research* 6, 1 (1981), 58–73.
- [23] Canice Prendergast. 2017. The Allocation of Food to Food Banks. (2017). Booth School of Business Working Paper.
- [24] Canice Prendergast. 2017. How Food Banks Use Markets to Feed the Poor. *Journal of Economic Perspectives* 31, 4 (2017), 145–62.
- [25] A. W. van der Vaart. 1998. *Asymptotic Statistics*. Cambridge University Press.
- [26] Hal R. Varian. 2007. Position auctions. *International Journal of Industrial Organization* 25, 6 (2007), 1163–1178.