Allocation for Social Good: Auditing Mechanisms for Utility Maximization

Taylor Lundy* Alexander Wei† Hu Fu‡ Scott Duke Kominers§
Kevin Leyton-Brown¶

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.

*University of British Columbia. tlundy@cs.ubc.ca
†Harvard University. weia@college.harvard.edu
‡University of British Columbia. hufu@cs.ubc.ca
§Harvard University. kominers@fas.harvard.edu
¶University of British Columbia. kevinlb@cs.ubc.ca

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. 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 [8] and Hartline and Roughgarden [20]; the mechanisms those authors derived naturally aim to minimize agents' payments because of payments' averse impact on total utility.[1] However, in charitable contexts, the imposition of even small payments may be unpalatable; moreover, as we have already described, ability to pay may be uncorrelated (or even inversely correlated) with need (see, e.g., Dworczak et al. [10]). We thus seek to 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 instead assume 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 agent not receiving an allocation would have been able to make use of one. 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 [8] and Hartline and Roughgarden [20]. 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 [8] and Hartline and Roughgarden [20] 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 [5, 15, 16]. de Clippel et al. [9] 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 11, 12]. 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 [7].

Another branch of work considers a tool more general than ours: *contingent payments*, in which payments 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 [19, 13]. There are also many examples of contingent payments being used in applications to both increase revenue and balance risk, including oil-lease auctions [18], ad auctions [27], and publishing rights [6]. The closest work in the contingent payment framework seeks to use contingent payments to increase social welfare. Most notably, Ma et al. [22] 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. [22] in some special cases (discussed in the sequel); however, a key difference is that Ma et al. [22] 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. [14] allocates goods in a finite time-horizon model. Two additional papers study the time-discounted infinite horizon model. Guo et al. [17] 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 [24, 25], 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 [21]).

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

Definition 3.1. *We say that $\mathcal{M} = (\boldsymbol{x}, \boldsymbol{p})$ 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 $\hat{\boldsymbol{t}} := (\hat{t}_1, \ldots, \hat{t}_N)$ to denote the resulting reported type profile.*
3. *The center makes an allocation $\boldsymbol{x}(\hat{\boldsymbol{t}}) = (x_1(\hat{\boldsymbol{t}}), \ldots, x_N(\hat{\boldsymbol{t}})) \in \mathbb{N}^n_{\geq 0}$, with $\sum_i x_i(\hat{\boldsymbol{t}}) \leq M$.*
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 $i \in \{1, \ldots, N\}$.*
6. *The center charges a payment $p_i(\hat{\boldsymbol{t}}, d_i^{\text{obs}}) \in \mathbb{R}^n_{\geq 0}$ 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 $p_i(\hat{\boldsymbol{t}}, d_i^{\text{obs}})$ as $p_i(\hat{\boldsymbol{t}})$. In general, the allocations $\boldsymbol{x}$ may be randomized. In this work, we discuss randomized mechanisms only in the unit-demand setting.

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 payments and utilities. Specifically, let

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

be the *interim payment given the realized demand.* The value $p_i(\hat{t}_i \mid d_i)$ represents the expected payment charged to agent i if they report $\hat{t}_i$ while all other agents report truthfully given their demand is d_i.

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

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

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

Definition 3.2. *An auditing mechanism $\mathcal{M}$ is* Bayesian–Nash incentive compatible *(BIC) if it makes honest reporting a Bayesian 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

Through classical auction theory, non-audited mechanisms are well-understood in the case where 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 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. Similarly, in the unit-demand setting, the allocation to any agent is either 0 or 1; as we consider in this setting randomized mechanisms (both auditing and non-auditing), the ex post allocation is a Bernoulli random variable, and the *interim allocation* of an agent i's type t_i can be represented as a real from $[0, 1]$, representing the probability with which the type is allocated a unit of good in the truthful equilibrium:

$$x_i(\hat{t}_i) \coloneqq \mathbf{E}_{\boldsymbol{t}_{-i} \sim \boldsymbol{G}_{-i}} \left[\mathbf{E} \left[x_i(\hat{t}_i, \boldsymbol{t}_{-i}) \right] \right].$$

Classical characterizations of single-parameter auctions carry over to non-auditing mechanisms:

Lemma 3.4 (Myerson [23])**.** *A non-audited mechanism $\mathcal{M} = (\boldsymbol{x}, \boldsymbol{p})$ 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*

$$p_i(\hat{t}_i) = \hat{t}_i x_i(\hat{t}_i) - \int_0^{\hat{t}_i} x_i(t) \, \mathrm{d}t + C \tag{MPI}$$

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 in the unit demand setting are 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 [20])**.** *Let $\mathcal{M} = (\boldsymbol{x}, \boldsymbol{p})$ be a BIC non-audited mechanism. Then the expected utility of agent i is*

$$\mathbf{E}_{\boldsymbol{t}} \left[t_i \cdot x_i(\boldsymbol{t}) - p_i(\boldsymbol{t}) \right] = \mathbf{E}_{\boldsymbol{t}} \left[\psi_i(t_i) \cdot x_i(\boldsymbol{t}) \right], \tag{1}$$

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

When the allocation is allowed to be randomized, in an auditing mechanism the ex post payment of agent i in general should be a function not only of the reported types $\hat{\boldsymbol{t}}$ and i's observed demand d_i^{obs}, but also of the realized allocation x_i. Conveniently, it is without loss of generality to only charge a payment when x_i is 1:

Proposition 3.6. *Given any (randomized) BIC auditing mechanism $\mathcal{M}$ for the unit-demand setting, there is another BIC auditing mechanism $\mathcal{M}'$ with the same ex post allocation rule as $\mathcal{M}$, such that in $\mathcal{M}'$ any agent makes a positive payment only when their allocation is* 1*.*

It is a fact that holds more generally, beyond unit-demand settings, that the mechanism only needs to charge when the allocation is high; we state this more formally and prove it in Appendix A. Proposition 3.6 comes as a consequence. In light of this, in the sequel we still refer to ex post payment as a function of only the reported types and the observed demand, written as $p_i(\boldsymbol{t}, d_i^{\text{obs}})$, with the understanding that the mechanism charges the agent only when the allocation is 1; when this occurs, agent i pays $p_i(\boldsymbol{t}, \min(d_i, 1)) / \mathbf{Pr}[x_i(\boldsymbol{t}) = 1]$.

4 Auditing Mechanisms for Unit-Demand Agents

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

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

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

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

Proof. We show that for any BIC auditing mechanism $\mathcal{M} = (\boldsymbol{x}, \boldsymbol{p})$, there is another BIC auditing mechanism $\mathcal{M}^* = (\boldsymbol{x}^*, \boldsymbol{p}^*)$ such that $p_i^*(\boldsymbol{t}, 1) = 0$ for all $i \in [N]$ and all $\boldsymbol{t}$, and $\mathcal{M}^*$ obtains the same social utility as $\mathcal{M}$.

For ease of notation, we let $y_i(t_i, \boldsymbol{t}_{-i})$ be $\mathbf{Pr}[x_i(t_i, \boldsymbol{t}_{-i}) = 1]$ and $y_i^*(t_i, \boldsymbol{t}_{-i})$ be $\mathbf{Pr}[x_i^*(t_i, \boldsymbol{t}_{-i}) = 1]$. In $\mathcal{M}$, agent i who has type t_i and reports $\hat{t}_i$, while the other agents report $\boldsymbol{t}_{-i}$, has ex post utility

$$t_i y_i(\hat{t}_i, \boldsymbol{t}_{-i}) - t_i p_i((\hat{t}_i, \boldsymbol{t}_{-i}), 1) - (1 - t_i) p_i((\hat{t}_i, \boldsymbol{t}_{-i}), 0).$$

We define new allocation and payment rules $\boldsymbol{p}^*$ and $\boldsymbol{x}^*$. For all i and $\boldsymbol{t}$, we set $p_i^*(\boldsymbol{t}, 1) = 0$, $p_i^*(\boldsymbol{t}, 0) = p_i(\boldsymbol{t}, 0)$ and $y_i^*(\boldsymbol{t}) = y_i(\boldsymbol{t}) - p_i(\boldsymbol{t}, 1)$. Note that if $p_i(\boldsymbol{t}, 1) = 0$, the allocation and payment for i in the new mechanism remains unchanged. First note that we have only decreased allocations in $\mathcal{M}^*$, so $\mathcal{M}^*$ is clearly feasible. We then show that $\mathcal{M}^*$ is BIC. In $\mathcal{M}^*$, agent i who has type t_i and reports $\hat{t}_i$, while the other agents report $\boldsymbol{t}_{-i}$, has ex post utility

$$\begin{aligned} & t_i(y_i^*(\hat{t}_i, \boldsymbol{t}_{-i}) - p_i^*((\hat{t}_i, \boldsymbol{t}_{-i}), 1)) - (1 - t_i) p_i^*((\hat{t}_i, \boldsymbol{t}_{-i}), 0) \\ =& t_i(y_i(\hat{t}_i, \boldsymbol{t}_{-i}) - p_i((\hat{t}_i, \boldsymbol{t}_{-i}), 1)) - (1 - t_i) p_i^*((\hat{t}_i, \boldsymbol{t}_{-i}), 0). \end{aligned}$$

Therefore the utility of any agent with any type for any report remains unchanged, and so $\mathcal{M}^*$ inherits incentive compatibility from $\mathcal{M}$. This completes the proof.

We remark that while the value of the ex post payment $p_i^*(\boldsymbol{t}, 0)$ remains unchanged from $p_i(\boldsymbol{t}, 0)$, the amount that is actually charged when the agent is allocated may increase, from $\frac{p_i(\boldsymbol{t},0)}{y_i(\boldsymbol{t})}$ to $\frac{p_i(\boldsymbol{t},0)}{y_i^*(\boldsymbol{t})}$, because the mechanisms only charge an agent when they are allocated a unit. (See Proposition 3.6 and the discussion that follows it.) □

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 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:*

(a) The interim allocation rule x_i is monotone non-decreasing.

(b) The expected payment for reporting $\hat{t}_i$ when the observed demand is 0 is

$$p_i(\hat{t}_i \mid 0) = C_i + \frac{\hat{t}_i \cdot x_i(\hat{t}_i)}{1 - \hat{t}_i} - \int_0^{\hat{t}_i} \frac{x(v)}{(1-v)^2} \,\mathrm{d}v, \tag{API}$$

where $C_i \geq 0$ is a constant.

As our objective is to maximize social utility, without loss of generality we may take the C_i in (API) to be 0.

Proof of Lemma 4.2. (⇐) We begin by proving that Bayesian incentive compatibility implies condition (a). Consider two types $t_i, t_i' \in T_i$. By BIC, type t_i has weakly higher utility for being truthful than for reporting t_i':

$$\begin{aligned} & t_i x_i(t_i) - (1-t_i) p_i(t_i \mid 0) \geq t_i x_i(t_i') - (1-t_i) p_i(t_i' \mid 0) \\ \Rightarrow \quad & \frac{t_i}{1-t_i}(x_i(t_i) - x_i(t_i')) \geq p_i(t_i \mid 0) - p_i(t_i' \mid 0). \end{aligned}$$

Similarly for t_i', we have

$$\begin{aligned} & t_i' x_i(t_i') - (1-t_i') p_i(t_i' \mid 0) \geq t_i' x_i(t_i) - (1-t_i') p_i(t_i \mid 0) \\ \Rightarrow \quad & \frac{t_i'}{1-t_i'}(x_i(t_i) - x_i(t_i')) \leq p_i(t_i \mid 0) - p_i(t_i' \mid 0). \end{aligned}$$

Comparing the two inequalities immediately gives

$$\frac{t_i'}{1-t_i'}(x_i(t_i) - x_i(t_i')) \leq \frac{t_i}{1-t_i}(x_i(t_i) - x_i(t_i')).$$

Since $\frac{z}{1-z}$ is an increasing function on $[0,1)$, we must have $x_i(t_i) \geq x_i(t_i')$ if $t_i \geq t_i'$. We now show that BIC implies condition (b). Fix some $t_i \in (0,1)$. By incentive compatibility, $u_i(t_i,t_i) \geq u_i(\hat{t}_i,t_i)$ for any $\hat{t}_i$, and first order condition gives

$$0 = \frac{\partial u_i(\hat{t}_i,t_i)}{\partial \hat{t}_i}\Big|_{\hat{t}_i=t_i} = t_i x_i'(t_i) - (1-t_i)p_i'(t_i \mid 0).$$

$$\Rightarrow p_i'(t_i \mid 0) = \frac{t_i}{1-t_i}x_i'(t_i).$$

The second equality follows from the Envelope Theorem [4, Chapter 6]. We can now integrate both sides from 0 to $\hat{t}_i$.

$$\int_0^{\hat{t}_i} p_i'(v \mid 0)\,\mathrm{d}v = \int_0^{\hat{t}_i} \frac{v}{1-v}x_i'(v)\,\mathrm{d}v.$$

$$\Rightarrow \quad p_i(\hat{t}_i \mid 0) - p_i(0 \mid 0) = \frac{\hat{t}_i x_i(\hat{t}_i)}{1-\hat{t}_i} - \int_0^{\hat{t}_i} \frac{x_i(v)}{(1-v)^2}\,\mathrm{d}v.$$

Letting $p_i(0 \mid 0)$ be C gives the identity in (API).

($\Rightarrow$) We now show that an auditing mechanism with a monotone interim allocation rule and payment identity (API) is BIC. We first show that any agent i cannot gain utility by increasing their bid by any $\delta > 0$. The proof for decreasing their bid by δ follows the same steps. For any $t_i \in [0, 1-\delta)$,

$$u(t_i+\delta, t_i) = t_i \cdot x_i(t_i+\delta) - (1-t_i)\left[\frac{(t_i+\delta)x_i(t_i+\delta)}{1-t_i-\delta} - \int_0^{t_i+\delta} \frac{x_i(v)}{(1-v)^2}\,\mathrm{d}v\right]$$

$$= \frac{-\delta x_i(t_i+\delta)}{1-t_i-\delta} + (1-t_i)\int_0^{t_i+\delta} \frac{x_i(v)}{(1-v)^2}\mathrm{d}v.$$

We now subtract this from the utility obtained from a truthful report:

$$u_i(t_i,t_i) - u_i(\hat{t}_i,t_i) = \frac{\delta x_i(t_i+\delta)}{1-t_i-\delta} - (1-t_i)\int_{t_i}^{t_i+\delta} \frac{x_i(v)}{(1-v)^2}\,\mathrm{d}v.$$

By the monotonicity of the allocation rule, it is easy to see that the integrand is upper bounded by $\frac{x_i(t_i+\delta)}{(1-v)^2}$, so

$$\frac{\delta x_i(t_i+\delta)}{1-t_i-\delta} - (1-t_i)\int_{t_i}^{t_i+\delta} \frac{x_i(v)}{(1-v)^2}\mathrm{d}v \geq \frac{\delta x_i(t_i+\delta)}{1-t_i-\delta} - (1-t_i)\left[\frac{x_i(t_i+\delta)}{1-t_i-\delta} - \frac{x_i(t_i+\delta)}{1-t_i}\right] = 0$$

This shows that agents have no incentive to overbid. A symmetric argument shows that underbidding is not profitable either. Therefore the mechanism is BIC. □

We now derive some immediate consequences of Lemma 4.2. We 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 [20])**.** *For $\alpha \geq 0$, the M-unit α-lottery allocates to agents with reported types at least α, and charges them each price α. If more than M agents have types at least α, M of them are selected uniformly at random as winners.*

Definition 4.4 (M-unit (α, β)-lottery – [20]). *For $\alpha \geq \beta \geq 0$, 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 a M-unit α-lottery.*
2. *If $|A| + |B| \leq M$, sell to the $|A| + |B|$ agents with the highest types at price β.*
3. *Otherwise, run a (M-$|A|$)-unit β-lottery for the agents in B, and allocate a unit to each agent in A at the price determined by Myerson's payment identity* (MPI).

Theorem 4.5 ([20]). *Given any type distributions $G_1, \cdots, G_n$, there is a mechanism that maximizes the expected social utility among all symmetric, BIC non-audited mechanisms, such that for every type profile $\boldsymbol{t}$, 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 $\boldsymbol{t}$, 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 corresponding 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 mechanism, if i is allocated an item (i.e., $t_i \geq \alpha$), a payment p is only charged when i does not use it, which occurs with probability $1 - t_i$. Incentive compatibility is guaranteed whenever p is high enough so that lower types have no incentive to misreport and get allocated an item; therefore p should be at least $\frac{\alpha}{1-t'_i}$ for any $t'_i < \alpha$; taking the minimum, *pay* is $\frac{\alpha}{1-\alpha}$. In expectation, agent i pays $\frac{1-t_i}{1-\alpha} \cdot \alpha$, which is lower than the price α in the non-audited mechanism, for any $t_i > \alpha$.[2]

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

[2] This example of modified second-price can also be derived via the approach in Ma et al. [22] when the center is only concerned with whether or not the item is consumed.

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 $\boldsymbol{t}$ is fixed in the theorem statement). The item is given with equal probability to each agent in A, so by Myerson's payment identity, $p_i = \frac{M}{|A|}\alpha$. Using our audited payment identity (API), we may calculate p_i^A:

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

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

4.2 Optimal Auditing Mechanism

The previous results demonstrate that whenever the symmetric 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 [20] showed that the optimal social utility mechanism is a *virtual utility maximizer*. In this section we derive a formulation of *audited* virtual utility 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 $\boldsymbol{x}$ achieves expected social utility*

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

This theorem lets us associate to each type t_i an *audited virtual utility* $x_i(t_i) \cdot \frac{(1-G_i(t_i))(1-\mathbf{E}[b|b\geq t_i])}{g_i(t_i)(1-t_i)^2}$, and maximizing social utility reduces 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 agent i' expected payment $\mathbf{E}_{t_i}[(1-t_i)p_i(t_i \mid 0)]$.

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

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

Proof. We begin by multiplying both sides of the payment identity (API) by $(1-t_i)$ and taking expectation over t_i:

$$
\mathbf{E}_{t_i}\left[(1-t_i)p_i(t_i \mid 0)\right] = \int_0^1 x_i(t_i)t_i g_i(t_i)\,\mathrm{d}t_i - \int_0^1\int_0^{t_i} \frac{x_i(v)}{(1-v)^2} g_i(t_i)(1-t_i)\,\mathrm{d}v\,\mathrm{d}t_i.
$$

We now swap the order of integrals in the second term.

$$\begin{aligned}
\mathbf{E}_{t_i}\left[(1-t_i)p_i(t_i \mid 0)\right] &= \int_0^1 x_i(t_i)t_i g_i(t_i)\,\mathrm{d}t_i - \int_0^1 \frac{x_i(v)}{(1-v)^2}\int_v^1 g_i(t_i)(1-t_i)\,\mathrm{d}t_i\,\mathrm{d}v \\
&= \int_0^1 x_i(t_i)t_i g_i(t_i)\,\mathrm{d}t_i - \int_0^1 \frac{x_i(v)}{(1-v)^2}\left(\int_v^1 g_i(t_i)\,\mathrm{d}t_i - \int_v^1 g_i(t_i)t_i\,\mathrm{d}t_i\right)\mathrm{d}v
\end{aligned}$$

The last integral computes the expectation of t_i when it lies in $[v, 1]$:

$$\int_v^1 g_i(t_i)t_i\,\mathrm{d}t_i = (1-G_i(v))\,\mathbf{E}_{b\sim G_i}\left[b \mid b \geq v\right].$$

Substituting this into our equation and solving the preceding integral we get

$$\mathbf{E}_{t_i}\left[(1-t_i)p_i(t_i \mid 0)\right] = \int_0^1 x_i(t_i)t_i g_i(t_i)\,\mathrm{d}t_i - \int_0^1 \frac{x_i(v)}{(1-v)^2}(1-G_i(v))(1-\mathbf{E}_{b\sim G_i}\left[b \mid b \geq v\right])\,\mathrm{d}v.$$

Finally we rename the variable v to t_i and factor out $x(t_i)f(t_i)$.

$$\begin{aligned}
\mathbf{E}_{t_i}\left[(1-t_i)p_i(t_i \mid 0)\right] &= \int_0^1 x_i(t_i)g_i(t_i)\left(t_i - \frac{(1-G_i(t_i))(1-\mathbf{E}_{b\sim G_i}[b \mid b \geq t_i])}{g_i(t_i)(1-t_i)^2}\right) \\
&= \mathbf{E}_{t_i}\left[x_i(t_i)\left(t_i - \frac{(1-G_i(t_i))(1-\mathbf{E}_{b\sim G_i}[b \mid b \geq t_i])}{g_i(t_i)(1-t_i)^2}\right)\right].
\end{aligned}$$

□

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 is

$$\sum_i \mathbf{E}_{t_i}\left[t_i \cdot x_i(t_i) - (1-t_i)p_i(t_i \mid 0)\right]. \tag{3}$$

Using Lemma 4.8, we can rewrite (3) as

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

□

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 social utility mechanism without auditing allocates the items uniformly at random without charging anyone (i.e., runs an M-unit 0-lottery), whereas the optimal social utility auditing mechanism is an M-unit α-lottery with $\mathbf{E}[\alpha] > 0$ and generates strictly more social utility.

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 $\lambda = 1$ truncated at 1 and renormalized. Namely, $g(t) = \frac{e^{-t}}{1-e^{-1}}$ and $G(t) = \frac{1-e^{-t}}{1-e^{-1}}$. Consider $G_1 = \cdots = G_n = G$, that is, all agents' types are drawn i.i.d. from G. The hazard rate for G is $\frac{e^{-t}}{e^{-t}-e^{-1}}$, which increases monotonically with t. Hartline and Roughgarden [20] showed that, in this case, the optimal social utility (non-auditing) mechanism allocates the item uniformly at random with no charge.

We now consider the optimal social utility auditing mechanism. From Theorem 4.7, we see that if $\frac{(1-G_i(t_i))(1-\mathbf{E}_{b\sim G_i}[b|b\geq t_i])}{g_i(t_i)(1-t_i)^2}$ is monotone increasing in t_i then the optimal social utility auditing mechanism should have interim allocation rule that solves

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

A calculation of the virtual utilities for distribution G shows that the audited virtual utility indeed strictly increases with t_i.[3] 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. On any reported type profile, this allocation rule is a 1-unit α-lottery, where α is second highest reported type. We also see that giving away the item for free is not a solution to the above maximization problem. Therefore, the auditing mechanism must achieve strictly higher utility in this setting. □

5 Beyond Unit Demand

When we go beyond unit-demand settings, the agents' preferences are multi-dimensional. Beyond the VCG mechanism, 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 increase social utility. We propose two variants of the VCG mechanism 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.

We first derive the VCG mechanism for our context. Recall that the VCG mechanism allocates items to maximize social welfare. Let us first see that its allocation rule in our setting 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 since $F_{t_i}(k)$ is a monotone increasing function, this marginal value decreases as k grows. Therefore, the problem boils down to 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 an agent i that is already allocated k items has marginal value $1 - F_{t_i}(k)$. Let $\boldsymbol{x}^{\mathbf{VCG}}$ denote the resulting allocation rule. The VCG mechanism then charges each agent i the externality that agent i causes. In this setting, the payment $p_i^{\mathrm{VCG}}(\boldsymbol{t})$ is the sum of the $(M - x_i^{\mathrm{VCG}}(\boldsymbol{t}) + 1)$-st through the M-th highest marginal utilities that do not belong to agent i.

We introduce a few notations to simplify the expressions. For $k = 1, 2, \ldots, M$, 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 $\nu_i^k(\boldsymbol{t}_{-i})$ 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 $\mu_i^k \geq \nu_i^k$ (up to tie-breaking), and the payment of agent i in the VCG mechanism is $p_i^{\mathrm{VCG}}(\boldsymbol{t}) = \sum_{k=0}^{x_i^{\mathrm{VCG}}(\boldsymbol{t})-1} \nu_i^k(\boldsymbol{t}_{-i})$.

[3] This calculation is simple algebra and is deferred to Appendix D, alongside a plot.

5.1 The Audited VCG (AVCG) Mechanism

In this section we 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*.

Definition 5.1. *The* audited VCG *(*AVCG*) mechanism has the same allocation rule as the VCG mechanism, and has payment rule*

$$p_i^{\mathrm{AVCG}}(\boldsymbol{t}, d_i^{\mathrm{obs}}) := \sum_{k=0}^{x_i^{\mathrm{VCG}}(\boldsymbol{t})-1} \frac{\mathbf{1}_{d_i^{\mathrm{obs}} \leq k}}{1-\nu_i^k(\boldsymbol{t}_{-i})} \nu_i^k(\boldsymbol{t}_{-i}) = \begin{cases} 0, & \text{if } d_i^{\mathrm{obs}} = x_i^{\mathrm{VCG}}(\boldsymbol{t}), \\ \sum_{k=d_i^{\mathrm{obs}}}^{x_i^{\mathrm{VCG}}(\boldsymbol{t})-1} \frac{\nu_i^k(\boldsymbol{t}_{-i})}{1-\nu_i^k(\boldsymbol{t}_{-i})}, & \text{otherwise,} \end{cases}$$

where $\mathbf{1}_{\mathscr{A}}$ is the indicator function for the event $\mathscr{A}$.

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 $\nu_i^k(\boldsymbol{t}_{-i})$, regardless of the 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 at least k units, but uses less than k units, the agent pays $\nu_i^k(\boldsymbol{t}_{-i})/(1-\nu_i^k(\boldsymbol{t}_{-i}))$ for the k-th unit; this value can be thought of as the VCG payment divided by a factor of $(1-\nu_i^k(\boldsymbol{t}_{-i}))$. Although this is larger than the VCG payment, *in expectation* a truthfully reporting agent pays no more than in the VCG mechanism, as we show formally in Theorem 5.2. (This is because this amount is charged only if the agent's realized demand is lower than k.) We show moreover that any non-truthful agent that gets extra allocations by misreporting should expect payments higher than in VCG; the intuition here is that such agents are punished more heavily for failing to use the 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 $\boldsymbol{t}_{-i}$, as

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

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 allocation and receives more items than they would under truthful reporting, then $\mu_i^k(t_i) \leq \nu_i^k(\hat{\boldsymbol{t}}_{-i})$ for each additional unit k they receive. It follows that $\frac{1-\mu_i^k(t_i)}{1-\nu_i^k(\hat{\boldsymbol{t}}_{-i})} \geq 1$, in which case the expected payment for the k-th unit is $\frac{1-\mu_i^k(t_i)}{1-\nu_i^k(\hat{\boldsymbol{t}}_{-i})}\nu_i^k(\hat{\boldsymbol{t}}_{-i}) \geq \mu_i^k(t_i)$. This shows that the agent incurs non-positive expected return for any unit of good allocated to them only due to misreporting information. In other words, the agent is weakly better off reporting truthfully.

On the other hand, if by misreporting, agent i changes the allocation and receives fewer items than they would 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, because the expected utility from any allocated unit under truthful reporting is non-negative. (Formally, for $k < x_i^{\mathrm{VCG}}(t_i, \boldsymbol{t}_{-i}), \mu_i^k(t_i) \geq \nu_i^k(\boldsymbol{t}_{-i}) \geq \frac{1-\mu_i^k(t_i)}{1-\nu_i^k(\boldsymbol{t}_{-i})}\nu_i^k(\boldsymbol{t}_{-i})$.) Furthermore, their expected payments for the units allocated to them remain unchanged from under truthful reporting, since the expected payments for agent i are determined entirely by the bids of the other agents and agent i's true type (but not $\hat{t}_i$, i's reported type). Thus agent i never benefits from misreporting to receive fewer items.

Agent i does not benefit from changing the allocation from that under truthful play, regardless of what the other agents do, so the AVCG mechanism is dominant strategy incentive compatible. □

Theorem 5.3. *In comparision to VCG, the AVCG mechanism satisfies the following:*

(a) For each agent i with type t_i, for any reported type profile $\boldsymbol{t}_{-i}$ of the other agents, the expected payment made by agent i with truthful bidding is at most the agent's VCG payment.

(b) For any type profile $\boldsymbol{t}$ with which agent i makes a strictly positive payment and obtains strictly positive utility in the VCG mechanism, the agent obtains strictly higher expected utility in the AVCG mechanism.

Proof. Notice that the condition $\mu_i^k(t_i) \geq \nu_i^k(\boldsymbol{t}_{-i})$ for allocating a k-th unit to agent i in the VCG allocation rule implies $\frac{1-\mu_i^k(t_i)}{1-\nu_i^k(\boldsymbol{t}_{-i})}\nu_i^k(\boldsymbol{t}_{-i}) \leq \nu_i^k(\boldsymbol{t}_{-i})$. Summing over k, we see that the expected AVCG payment is always at most the VCG payment. Furthermore, if for any i and k, $\mu_i^k(t_i) > \nu_i^k(\boldsymbol{t}_{-i}) > 0$, then the inequality is strict; this is exactly the condition under which VCG charges a positive payment and generates positive utility. □

Our construction of the AVCG mechanism shows that auditing generally 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 welfare-maximizing allocation rule and dominant strategy incentive compatibility, without making any assumption 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 each agent's type space admits an ordering by stochastic dominance.

Definition 5.4. *An agent i's type space admit a* First-order Stochastic Dominance Ordering (FSDO) *if there exists a total ordering "$\geq$" on T_i such that $t_i \geq t_i'$ if and only if $\mu_i^k(t_i) \geq \mu_i^k(t_i')$ for any demand k.*

We define a mapping that allows us to relate the marginal values of opponents with the marginal values of each agent's other potential types. For any opponent's marginal value ν, and $j, k \in \mathbb{N}_{\geq 0}$, let $\psi_i(\nu, k, j)$ be $\sup_{\{t_i : \mu_i^j(t_i) \leq \nu\}} \mu_i^k(t_i)$. Intuitively, if the supremum is obtained by type t_i, this type can be thought of as the highest type that i can have without being allocated at least $(j+1)$ units when an opponent with marginal value ν is not allocated the corresponding unit. $\psi_i(\nu, k, j)$ is then this type t_i's marginal value for the $(k+1)$-st item.

Definition 5.5. *For agents whose type spaces each admit a first-order stochastic dominance ordering, the* SD-AVCG *mechanism has the same allocation rule as the VCG mechanism, and has the following payment rule. Let*

$$d_i(t_i, \boldsymbol{t}_{-i}, j) \coloneqq \operatorname*{arg\,min}_{k < x_i^{\mathrm{VCG}}(t_i, \boldsymbol{t}_{-i})} \left[\frac{1 - \mu_i^k(t_i)}{1 - \psi_i(\nu_i^j(\boldsymbol{t}_{-i}), k, j)} \right].$$

The (ex post) payment rule of SD-AVCG is

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

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

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

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

(a) For each agent i with type t_i, for any reported type profile $\boldsymbol{t}_{-i}$ of the other agents, the expected payment made by agent i with truthful bidding is at most the agent's AVCG payment.

(b) For any type profile $\boldsymbol{t}$ with which agent i makes a strictly positive payment and obtains strictly positive utility in the VCG mechanism, the agent obtains strictly higher expected utility in the SD-AVCG mechanism than in the VCG mechanism.

Proof. We first prove dominant-strategy incentive compatibility. If agent i with type t_i misreports and gets allocated more items than with truthful reporting, for each additional unit he receives he is charged an additional payment that is comprised of two terms. Recall that the expected payment made by the agent for their $(\ell+1)$-st allocated good is

$$\psi_i(\nu_i^\ell(\boldsymbol{t}_{-i}), \ell, \ell) \min_{k \leq \ell} \left[\frac{1-\mu_i^k(t_i)}{1-\psi_i(\nu_i^\ell(\boldsymbol{t}_{-i}), k, \ell)}\right].$$

If the agent would not be allocated this $(\ell+1)$-st good with truthful reporting, we show this payment is larger than the agent's marginal gain for the item, i.e.,

$$\mu_i^\ell(t_i) \leq \psi_i(\nu_i^\ell(\boldsymbol{t}_{-i}), \ell, \ell) \min_{k \leq \ell} \left[\frac{1-\mu_i^k(t_i)}{1-\psi_i(\nu_i^\ell(\boldsymbol{t}_{-i}), k, \ell)}\right].$$

It is easy to see that $\mu_i^\ell(t_i) \leq \psi_i(\nu_i^\ell(\boldsymbol{t}_{-i}), \ell, \ell)$ since agent i is allocated at most ℓ units with truthful reporting and $\psi_i(\nu_i^j(\boldsymbol{t}_{-i}), \ell, \ell)$ is by definition the maximum marginal value for the $(\ell+1)$-st item over the agent's types that are allocated at most ℓ items (when the opponents report $\boldsymbol{t}_{-i}$). It remains to show that the factor $\min_{k \leq \ell}\left[\frac{1-\mu_i^k(t_i)}{1-\psi_i(\nu_i^\ell(\boldsymbol{t}_{-i}),k,\ell)}\right]$ is at most one. We know that $\forall k \leq \ell$, $\mu_i^k(t_i) \leq \psi_i(\nu_i^\ell(\boldsymbol{t}_{-i}), k, \ell)$ by definition of $\psi_i(\nu_i^\ell(\boldsymbol{t}_{-i}), k, \ell)$. Therefore, for all $k \leq \ell$, $\frac{1-\mu_i^k(t_i)}{1-\psi_i(\nu_i^\ell(\boldsymbol{t}_{-i}),k,\ell)}$ is at least 1. Combining these two terms makes it clear the agent has no incentive to deviate to receive extra items. Similarly if any agent misreports and gets allocated fewer items than with truthful reporting, their utility is weakly worse.

For part (a), we show that the payment in SD-AVCG for each item is upper bounded by the corresponding payment in AVCG. Note that for any opponents' type profile $\boldsymbol{t}_{-i}$, $\psi_i(\nu_i^j(\boldsymbol{t}_{-i}), j, j) \leq \nu_i^j(\boldsymbol{t}_{-i})$. Therefore the first term in the payment for the $(j+1)$-st item in SD-AVCG is no larger than in AVCG. For the second term, we are comparing $\min_{k < x_i^{\text{VCG}}(\hat{t}_i, \boldsymbol{t}_{-i})}\left[\frac{1-\mu_i^k(\hat{t}_i)}{1-\psi_i(\nu_i^j(\boldsymbol{t}_{-i}),k,j)}\right]$ in SD-AVCG with $\frac{1-\mu_i^j(t_i)}{1-\nu_i^j(\boldsymbol{t}_{-i})}$ in AVCG. Again, we have

$$\min_{k < x_i^{\text{VCG}}(t_i, \boldsymbol{t}_{-i})} \left[\frac{1-\mu_i^k(\hat{t}_i)}{1-\psi_i(\nu_i^j(\boldsymbol{t}_{-i}), k, j)}\right] \leq \frac{1-\mu_i^k(t_i)}{1-\psi_i(\nu_i^j(\boldsymbol{t}_{-i}), j, j)} \leq \frac{1-\mu_i^j(t_i)}{1-\nu_i^j(\boldsymbol{t}_{-i})}.$$

Therefore, the SD-AVCG payment is at most the AVCG payment.

Part (b) follows from combining Theorem 5.3 with part (a). □

A simple example where 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 for settings with repeated interactions. Debt mechanisms give a way to extend any static auditing mechanism into a repeated auditing 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.

Our reduction relies on a variant of the classic observation that repetition allows us to use future allocative penalties to disincentivize 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 where imposing any sort of monetary transfer may be inappropriate.

6.1 Setup

A *debt mechanism* is a non-monetary repeated mechanism that runs for infinitely many rounds. It is described by a single-round quasilinear mechanism $\mathcal{M} = (\boldsymbol{x}, \boldsymbol{p})$, an *allocation length* ℓ, and a *debt rate* r_i for each agent i. We denote by $\boldsymbol{r} \in \mathbb{R}_+^N$ the vector of debt rates, and the debt mechanism is denoted as $\mathcal{M}_D = (\mathcal{M}, \boldsymbol{r}, \ell)$. For debt mechanisms, we again make the assumption of non-negative payments for $\mathcal{M}$, i.e, $p_i \geq 0$ always.[4]

A debt mechanism records each agent's total debt to the mechanism over time, and in any given round, 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.

[4]Note 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 to settings with less frequent (or stochastic) auditing.

For each agent i, the rounds are divided into *intervals*. Each interval starts with ℓ guaranteed consecutive allocation rounds (called an *allocation interval*), followed by a series of punishment rounds, whose length (possibly 0, and possibly randomized) is determined by the debt accumulated in the ℓ allocation rounds and the agent's debt rate r_i, in the following manner: after the ℓ allocation rounds, before each round starts, if the agent's debt is at least r_i, their debt is reduced by r_i and the agent enters a punishment round; if the agent's debt is $b < r_i$, the debt is reset to 0 and with probability b/r_i the agent enters a pubnishment round, and with probability $1 - \frac{b}{r_i}$ a new interval starts (with an allocation round). For example, if $\ell = 2$ and $r_i = 3$ and agent i accumulates a debt of 5 over the first two rounds, round 3 is a punishment round, and with probability $\frac{2}{3}$, round 4 is a punishment round, in which case a new interval starts in round 5, and with probability $\frac{1}{3}$, a new interval starts in round 4.

We define a *round indicator* 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.

Strategy Space. We now formalize the strategy space available to the agents. We begin by considering the case with a single agent. For ease of notation, we drop the subscript referring to the agent number. In 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 number of rounds 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 $H^j \coloneqq \{(a^1, \ldots, a_i^{j-1}), (x(\hat{t}^1), \ldots, x(\hat{t}^{j-1})), (p(\hat{t}^1), \ldots, p(\hat{t}^{j-1}))\} \cup \{a^j\}$, where $\hat{t}^k$ is the type reported to the mechanism in round k.[5] Let $\mathcal{H}$ be the space of all potential histories. A strategy for an agent is then a map $\sigma : \mathcal{H} \times \mathbb{Z}_{>0} \times T \to T$, mapping tuples of $(\text{history}, \text{round number}, \text{type})$ to a reported type.

Values. By playing a strategy σ in a debt mechanism $\mathcal{M}_D = ((x, p), \ell, r)$, the agent receives value in the allocation rounds, accumulated additively over time. Formally, the value the agent receives in an allocation round j is determined by t^j, their type in round j, and their allocation x in that round. Let $V(t^j, x) \coloneqq \mathbf{E}_{d \sim t^j}[\min(d, x)]$ denote this value. Let H^j denote the history up to round j when the agent plays strategy σ, then the expected value accumulated after round j is

$$\sum_{j=1}^{n} \mathbf{E}\left[a^j \cdot V(t^j, x(\sigma(H^j, j, t^j)))\right].$$

We assume the agent's preference over the strategies is determined by the following overtaking criterion.

Definition 6.1 (Overtaking Criterion)**.** *For two strategies σ and $\hat{\sigma}$, for each round j, let a^j and $\hat{a}^j$ be the round indicators for round j when the agent uses strategy σ and $\hat{\sigma}$, respectively, and similarly let H^j and$\hat{H}^j$ denote the history up to round j under σ and $\hat{\sigma}$, respectively. σ is* weakly preferred *to $\hat{\sigma}$ by the agent if and only if*

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

[5]When the allocation and payment rules are randomized, the history stores the realized allocations and payments, rather than their expectations.

We remark that the preference defined by the overtaking criterion gives rise only to a partial order, instead of a total order, over the space of strategies. There may be strategies σ and $\hat{\sigma}$ such that

$$\liminf_{n\to\infty} \sum_{j=1}^{n} \left(\mathbf{E}\left[a^j \cdot V(t^j, x(\sigma(H^j, j, t^j)))\right] - \mathbf{E}\left[\hat{a}^j \cdot V(t^j, x(\hat{\sigma}(\hat{H}^j, j, t^j)))\right] \right) < 0,$$

$$\liminf_{n\to\infty} \sum_{j=1}^{n} \left(\mathbf{E}\left[\hat{a}^j \cdot V(t^j, x(\hat{\sigma}(\hat{H}^j, j, t^j)))\right] - \mathbf{E}\left[a^j \cdot V(t^j, x(\sigma(H^j, j, t^j)))\right] \right) < 0,$$

in which case σ and $\hat{\sigma}$ are incomparable.

Optimal Strategies. A strategy σ is *optimal* if any other strategy $\hat{\sigma}$ is not strictly preferred over σ:

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

Roadmap for the Rest of Section 6. The goal of this section is to show that any truthful single-round mechanism with monetary payments can be implemented by a debt mechanism in which reporting types truthfully in every allocation round is an optimal strategy for all the agents and yields a time-average utility equal to the agent's ex ante utility in the single-round mechanism. Towards this goal, we first focus on single-agent mechanisms in Section 6.2, Section 6.3 and Section 6.4; conclusions we obtain are then quickly generalized to multi-agent mechanisms in Section 6.5.[6] In Section 6.2 and Section 6.3, we focus on debt mechanisms with allocation length $\ell = 1$. In Section 6.2, we show that any debt mechanism admits some optimal strategy that is constant over time and oblivious of the history. We call such a strategy *stationary* and define it formally in Section 6.2. In Section 6.3, we define a class of truthful debt mechanisms, built on single-round truthful monetary mechanisms. We show that the time-average values of stationary strategies in these mechanisms converge; for the truthful strategy, this value converges to the ex ante utility in the single-round mechanism, and that of all other stationary strategies converges to a utility weakly worse. We then generalize the results in Section 6.2 and Section 6.3 to debt mechanisms with $\ell > 1$ in Section 6.4, making use of the time-average argument developed in Section 6.3.

6.2 Stationary Strategies

Recall that we focus on single-agent mechanisms till Section 6.5, and omit all subscripts for agent identities.

A strategy is said to be *stationary* if it does not depend on time or history. Formally,

Definition 6.2. *A strategy σ is* stationary *if and only if, for any type $t \in T$, round numbers $k, k' \in \mathbb{Z}_{>0}$, and histories H and H' up to rounds k and k', respectively, we have $\sigma(H, k, t) = \sigma(H', k', t)$.*

Abusing notations slightly, we write a stationary strategy σ as a mapping from T to T, with $\sigma(t)$ denoting the type reported to the debt mechanism in any round by the agent with type t in

[6] Proper discussion of the multi-agent case, though simple, needs conclusions from Section 6.2 to Section 6.4, so we defer formal definitions till Section 6.5.

that round. We now show that, for any debt mechanism with allocation length $\ell = 1$, there exists an optimal strategy that is stationary. The same is true for general allocation lengths; we defer its discussion to Section 6.4.

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

$$\liminf_{n\to\infty} \left(\sum_{j=1}^{n} \mathbf{E}\left[a^j \cdot V(t^j, x(\hat{\sigma}(\hat{H}^j, j, t^j)))\right] - \sum_{j=1}^{n} \mathbf{E}\left[a^j \cdot V(t^j, x(\sigma(t^j)))\right] \right) \leq 0,$$

where $\hat{H}^j$ is history up to round j if the agent plays strategy $\hat{\sigma}$.

Proof Sketch. We first show the existence of an optimal strategy whose mapping at any round k is independent of the history. We then use this to show that, for an optimal strategy, its sequence of strategies after round k' can be replaced by its sequence of strategies starting at another time k, without diminishing the expected cumulative values. We relegate the full proof to Appendix C. □

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.

6.3 Truthful Mechanisms Are Optimal

We now introduce the notion of *truthful debt mechanisms.* For a single agent, a debt mechanism $\mathcal{M}_D = (\mathcal{M}, r, \ell)$ is *truthful* if its single-round mechanism $\mathcal{M} = (x, p)$ satisfies (Bayesian) incentive compatibility constraints, with strictly positive ex ante utility for a truthful agent, and the debt rate r is $\mathbf{E}_{t\sim G}[V(t, x(t)) - p(t)] > 0$. (Recall that $V(t, x) = \mathbf{E}_{d\sim t}[\min(d, x)]$ is the expected value realized by the agent with type t when receiving allocation x.) We show that, in a truthful debt mechanism, there is an optimal strategy which in every round reports the true type, i.e., is stationary and truthful.

By the existence of an optimal stationary strategy (Theorem 6.3), we only need to compare being truthful with other stationary strategies. We relate the time average utility of a stationary strategy to the (ex ante) utility in the underlying single-round mechanism $\mathcal{M}$. In particular, we show that the utility of a stationary strategy, averaged over the first N rounds, converges almost surely as N goes to infinity. We show that for the truthful stationary strategy, this average converges to the ex ante (expected) utility for being truthful in the single-round mechanism, which is no worse than any other stationary strategy. By overtaking criterion, this immediately shows that the truthful strategy is an optimal strategy.

We will use a version of strong law of large numbers that allows the number of random variables to be random. For completeness we provide a proof in Appendix C.

Lemma 6.4. *Let $\{X_n, n \geq 1\}$ be an i.i.d. sequence of random variables such that $\mathbf{E}[|X_1|] < \infty$, and $\{I_n, n \geq 1\}$ another sequence of random variables which are allowed to be dependent on $\{X_n\}$ and such that I_n takes positive integer values and weakly increases with n, and almost surely goes to infinity with n. Then $\bar{X}_n := \frac{1}{I_n}\sum_{i=1}^{I_n} X_i$ converges almost surely to $\mathbf{E}[X_1]$.*

Recall that a^j is the indicator variable for the j-th round being an allocation round.

Lemma 6.5. *When the agent plays a stationary strategy σ, $\frac{1}{n}\sum_{j=1}^{n} a^j$ converges almost surely to $\frac{r}{r+\mathbf{E}_t[p(\sigma(t))]}$.*

Proof. For each round n, let I_n be the number of intervals that have started up to round n. (All but at most one of the I_n intervals are finished by round n.) As the type space is finite, the payment is bounded by some $p_{\max} > 0$, and the length of any interval is bounded by $\ell + \lceil \frac{p_{\max}}{r} \rceil$. Therefore, almost surely, I_n goes to infinity with n.

For the i-th interval starting in round π_i, let P_i be the number of punishment rounds of that interval, then $\mathbf{E}[P_i] = \mathbf{E}[\sum_{j=\pi_i}^{\pi_i+\ell-1} p(\sigma(t^j))]/r = \ell\, \mathbf{E}_t[p(\sigma(t))]/r$. Over the first n rounds, the number of allocation rounds is $\ell(I_n - 1) + \min(\ell, n - \pi_{I_n} + 1)$, and the number of punishment rounds is $\sum_{i=1}^{I_n-1} P_i + \max(0, n - (\pi_{I_n} + \ell - 1))$.

$$\frac{1}{n}\sum_{j=1}^{n} a^j = \frac{\ell(I_n - 1) + \min(\ell, n - \pi_{I_n} + 1)}{\ell(I_n - 1) + \sum_{i=1}^{I_n-1} P_i + n - \pi_{I_n} + 1} = \frac{\ell + \frac{\min(\ell, n-\pi_{I_n}+1)}{I_n-1}}{\ell + \frac{1}{I_n-1}\sum_{i=1}^{I_n-1} P_i + \frac{n-\pi_{I_n}+1}{I_n-1}}$$

Note that ℓ and $n - \pi_{I_n} + 1$ are both upper bounded by the maximum interval length, as I_n grows large with n, we have, almost surely,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} a^j = \frac{\ell}{\ell + \frac{1}{I_n-1}\sum_{i=1}^{I_n-1} P_i}.$$

Note that $\{P_i, i \geq 1\}$ is an i.i.d. sequence of random variables with $\mathbf{E}[|P_1|] \leq \ell p_{\max} < \infty$. Applying Lemma 6.4, we have $\frac{1}{I_n-1}\sum_{i=1}^{I_n-1} P_i$ converges almost surely to $\mathbf{E}[P_1] = \ell\, \mathbf{E}_t[p(\sigma(t))]/r$. Now by the Continuous Mapping Theorem [see, e.g. 26, Chapter 2], $\frac{1}{n}\sum_{j=1}^{n} a^j$ converges almost surely to $\frac{\ell}{\ell + \ell\, \mathbf{E}_t[p(\sigma(t))]/r} = \frac{r}{r + \mathbf{E}_t[p(\sigma(t))]}$.

□

Theorem 6.6. *In any truthful debt mechanism, reporting truthfully is an optimal strategy.*

Proof. By Theorem 6.3, we need only to focus on stationary strategies. Consider any stationary strategy σ, let $Y := \mathbf{E}[V(t, x(\sigma(t)))]$ be the expected value of playing σ in any allocation round. Let a^j be the indicator variable for round j being an allocation round when the agent plays σ. The expected value in round j is then $Y \cdot a^j$, since a^j and the value from the j-th round (if it were an allocation round) are independent random variables. (This is because a^j is determined entirely by the outcomes of previous rounds, whereas the value of the agent in round j, if it were an allocation round, is solely a function of the agent's type, which is drawn independently from the previous types.) Then the average expected utility after n rounds, $\frac{1}{n}\sum_{j=1}^{n} Y \cdot a^j$, by Lemma 6.5, converges almost surely to $Y \cdot \frac{r}{r + \mathbf{E}_t[p(\sigma(t))]} = \frac{\mathbf{E}[V(t, x(\sigma(t)))]}{r + \mathbf{E}_t[p(\sigma(t))]} \cdot r$.

Recall that in a truthful debt mechanism, we set r to be $\mathbf{E}[V(t) - p(t)]$. Now since $\mathcal{M}$ is BIC, the ratio

$$\frac{\mathbf{E}[V(t, x(\sigma(t)))]}{r + \mathbf{E}_t[p(\sigma(t))]} = \frac{\mathbf{E}[V(t, x(\sigma(t)))]}{\mathbf{E}[V(t, x(t)) - p(t)] + \mathbf{E}[p(\sigma(t))]} \leq 1$$

for any stationary strategy σ, with equality attained by the truthful strategy. As a consequence, for any $\epsilon > 0$, the expected utility from being truthful cannot be smaller by ϵ infinitely often than the utility from any other strategy. By definition this means being truthful is an optimal strategy in a truthful debt mechanism. □

The following corollary immediately follows from the proof of Theorem 6.6:

Corollary 6.7. *For any truthful debt mechanism, the time-average utility of the truthful strategy converges almost surely to the social utility of the single-round mechanism.*

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

$$\mathbf{E}_{\boldsymbol{t}}\left[V(t, x(\sigma(t))) - p(\sigma(t))\right] \leq \mathbf{E}\left[V(t, x(t)) - p(t)\right] \tag{4}$$

for any strategy σ. The next proposition shows that (4) is in fact equivalent to satisfying the BIC constraints on all sets of types which appear with non-zero probability.

Proposition 6.8. *A mechanism $\mathcal{M}$ satisfies BIC constraints if* (4) *holds for all mappings σ.*

The proof of Proposition 6.8 is a simple application of IC constraints and is shown in Appendix C.

6.4 Debt Mechanisms with Allocation Length $\ell > 1$

6.5 Multi-Agent Debt Mechanisms

Our discussion so far in this section addressed mechanisms for incentivizing individual agents to reveal their types truthfully. For multiple agents, a debt mechanism $\mathcal{M}_D = (\mathcal{M}, \boldsymbol{r}, \ell)$ is *truthful* if its single-round mechanism $\mathcal{M} = (\boldsymbol{x}, \boldsymbol{p})$ satisfies Bayesian incentive compatibility constraints, and for each agent i, the debt rate $r_i = \mathbf{E}_{\boldsymbol{t}}[V_i(t_i, x_i(\boldsymbol{t})) - p_i(\boldsymbol{t})]$. 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 indifferent as to whether their opponents are in punishment or allocation rounds (because agents' priors over possible opponent types remain unchanged).

7 Conclusions and Future Work

We have explored how auditing can help us maximize social welfare, as measured by total utility; 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 motivating 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. Multi-agent mechanism design without money. Working Paper, 2017.

[2] Elchanan Ben-Porath, Eddie Dekel, and Barton L. Lipman. Optimal allocation with costly verification. *American Economic Review*, 104(12):3779–3813, 2014.

[3] Jeremiah Blocki, Nicolas Christin, Anupam Datta, Ariel D. Procaccia, and Arunesh Sinha. Audit games. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*, pages 41–47, 2013.

[4] Michael Carter. *Foundations of Mathematical Economics*. Foundations of Mathematical Economics. MIT Press, 2001. ISBN 9780262531924.

[5] Ruggiero Cavallo. 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*, pages 882–889. ACM, 2006.

[6] Richard E. Caves. Contracts between art and commerce. *Journal of Economic Perspectives*, 17 (2):73–83, 2003.

[7] Sofia Ceppi, Ian Kash, and Rafael Frongillo. Partial verification as a substitute for money. *arXiv:1812.07312*, 2018.

[8] Surajeet Chakravarty and Todd Kaplan. Optimal allocation without transfer payments. *Games and Economic Behavior*, 77(1):1–20, 2013.

[9] Geoffroy de Clippel, Victor Naroditskiy, Maria Polukarov, Amy Greenwald, and Nicholas R. Jennings. Destroy to save. *Games and Economic Behavior*, 86:392–404, 2014.

[10] Piotr Dworczak, Scott Duke Kominers, and Mohammad Akbarpour. Redistribution through markets. Becker Friedman Institute Working Paper, 2019.

[11] Albin Erlanson and Andreas Kleiner. Costly verification in collective decisions. Working Paper, 2015.

[12] Albin Erlanson and Andreas Kleiner. A note on optimal allocation with costly verification. Working Paper, 2018.

[13] Thomas Giebe and Elmar Wolfstetter. License auctions with royalty contracts for (winners and) losers. *Games and Economic Behavior*, 63(1):91–106, 2008.

[14] Artur Gorokh, Siddhartha Banerjee, and Krishnamurthy Iyer. 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*, pages 563–564, 2017.

[15] Mingyu Guo and Vincent Conitzer. Worst-case optimal redistribution of VCG payments in multi-unit auctions. *Games and Economic Behavior*, 67(1):69–98, 2009.

[16] Mingyu Guo and Vincent Conitzer. Optimal-in-expectation redistribution mechanisms. *Artificial Intelligence*, 174(5-6):363–381, 2010.

[17] Mingyu Guo, Vincent Conitzer, and Daniel M. Reeves. Competitive repeated allocation without payments. In *Proceedings of the 5th International Workshop on Internet and Network Economics, Rome, Italy, December 14-18, 2009*, pages 244–255, 2009.

[18] Philip Haile, Kenneth Hendricks, and Robert Porter. Recent us offshore oil and gas lease bidding: A progress report. *International Journal of Industrial Organization*, 28(4):390–396, 2010.

[19] Robert G. Hansen. Auctions with contingent payments. *American Economic Review*, 75(4): 862–865, 1985.

[20] Jason D. Hartline and Tim Roughgarden. 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*, pages 75–84, 2008.

[21] Scott Duke Kominers and Alan Lam. Feeding America (A) and (B). Harvard Business School Case 818-130, Supplement 818-131, and Teaching Note 918-082, 2018–2019.

[22] Hongyao Ma, Reshef Meir, David C. Parkes, and James Zou. Contingent payment mechanisms to maximize resource utilization. *arXiv:1607.06511*, 2016.

[23] Roger Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

[24] Canice Prendergast. The allocation of food to food banks. Booth School of Business Working Paper, 2017.

[25] Canice Prendergast. How food banks use markets to feed the poor. *Journal of Economic Perspectives*, 31(4):145–62, 2017.

[26] A. W. van der Vaart. *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998.

[27] Hal R. Varian. Position auctions. *International Journal of Industrial Organization*, 25(6): 1163–1178, 2007.

A Proofs from Section 3

In this appendix we show that, when an auditing mechanism's allocation is allowed to be a lottery over natural numbers, it is without loss of generality to only charge a payment when an agent receives the highest allocation from the support of the lottery. In this paper we make use of randomized allocations only in the unit demand setting, but the proof we give here holds for the more general auditing mechanism outlined in Definition 3.1.

For any agent i and type profile $\boldsymbol{t}$, consider any ex post allocation $x_i(\boldsymbol{t})$ which is a random variable taking on values from $[M] = 0, 1, \ldots, M$. For $r \in [M]$, let $y_i^r(\boldsymbol{t})$ be $\mathbf{Pr}[x_i(\boldsymbol{t}) = r]$, the probability with which the allocation to i is r. As we discussed in Section 3, in general the ex post payment for agent i may be allowed to depend on the realization of $x_i(\boldsymbol{t})$; with a slight abuse of notation, we write such a payment rule as $p_i(\boldsymbol{t}, \min(x_i(\boldsymbol{t}), d_i), x_i(\boldsymbol{t}))$.

Theorem A.1. *For an agent i, type profile $\boldsymbol{t}$, and allocation lottery $x_i(\boldsymbol{t})$, let $r_{\max}$ be the maximum value in the support of $x_i(\boldsymbol{t})$. For any payment rule $p_i : \boldsymbol{T} \times \mathbb{N} \times \mathbb{N} \to \mathbb{R}_+$, there exists payment rule p_i^* such that*

$$\forall d_i, r \neq r_{\max}, \quad p_i^*(\boldsymbol{t}, \min(r, d_i), r) = 0;$$
$$\mathbf{E}_{r \sim x_i(\boldsymbol{t}), d_i \sim t_i}\left[p_i^*(\boldsymbol{t}, \min(r, d_i), r)\right] = \mathbf{E}_{r \sim x_i(\boldsymbol{t}), d_i \sim t_i}\left[p_i(\boldsymbol{t}, \min(r, d_i), r)\right].$$

Proof. Consider two allocations r_1, r_2 drawn from $x_i(\boldsymbol{t})$, with $r_1 > r_2$. We show there is a payment rule with the same expected payment as p_i but never charges when the allocation is r_2. Under payment rule p_i, the contributions to the expected payment when the allocation is r_1 and r_2 are respectively

$$y_i^{r_1}(\boldsymbol{t})\, \mathbf{E}_{d_i}\left[p_i(\boldsymbol{t}, \min(d_i, r_1), r_1)\right] = y_i^{r_1}(\boldsymbol{t}) \left(\mathbf{Pr}\left[d_i \geq r_1\right] p_i(\boldsymbol{t}, r_1, r_1) + \sum_{r=0}^{r_1 - 1} \mathbf{Pr}\left[d_i = r\right] p_i(\boldsymbol{t}, r, r_1) \right);$$
$$y_i^{r_2}(\boldsymbol{t})\, \mathbf{E}_{d_i}\left[p_i(\boldsymbol{t}, \min(d_i, r_2), r_2)\right] = y_i^{r_2}(\boldsymbol{t}) \left(\mathbf{Pr}\left[d_i \geq r_2\right] p_i(\boldsymbol{t}, r_2, r_2) + \sum_{r=0}^{r_2 - 1} \mathbf{Pr}\left[d_i = r\right] p_i(\boldsymbol{t}, r, r_2) \right).$$

Define ex post payment rule p_i^*: for each $r = 0, 1, \ldots, r_2 - 1$, set

$$p_i^*(\boldsymbol{t}, r, r_1) = p_i(\boldsymbol{t}, r, r_1) + \frac{y_i^{r_2}(\boldsymbol{t})}{y_i^{r_1}(\boldsymbol{t})} p_i(\boldsymbol{t}, r, r_2), \quad p_i^*(\boldsymbol{t}, r, r_2) = 0;$$

for $r = r_2, r_2 + 1, \ldots, r_1$, set

$$p_i^*(\boldsymbol{t}, r, r_1) = p_i(\boldsymbol{t}, r, r_1) + \frac{y_i^{r_2}(\boldsymbol{t})}{y_i^{r_1}(\boldsymbol{t})} p_i(\boldsymbol{t}, r_2, r_2), \quad p_i^*(\boldsymbol{t}, r, r_2) = 0;$$

and for all other values of r, r', $r \leq r'$, $p_i^*(\boldsymbol{t}, r, r') = p_i(\boldsymbol{t}, r, r')$. It is straightforward to check that the expected payment under p_i^* remains the same as under p_i, but $p_i^*(\boldsymbol{t}, \min(d_i, r_2), r_2) = 0$ for all realized demand d_i. Repeating this procedure eventually yields a payment rule p_i^* in the statement of the theorem.

□

B Proofs from Section 4

Theorem 4.6. *For any (α, β)-lottery and truthfully reported types $\boldsymbol{t}$, 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 corresponding audited mechanism generates strictly higher utility.

Proofs for the first two cases were given in Section 4. Here we give the proof for case 3 of the theorem.

Proof. Finding the difference in utility for agents with types in set B is identical to the proof of Theorem 4.6, so we focus on agents in the set A. Let p_i be the payment determined by the non-auditing mechanism and p_i^A the payment determined by the auditing mechanism for agents in A. For ease of notation let $x_\alpha \coloneqq 1$ and $x_\beta \coloneqq \frac{M - |A|}{|B|}$. We first compute the expected payment in the non-auditing auction as

$$p_i(t_i) = t_i \cdot x_\alpha - \int_\alpha^{t_i} x_\alpha \, \mathrm{d}v - \int_\beta^\alpha x_\beta \, \mathrm{d}v = \alpha \cdot x_\alpha - \alpha \cdot x_\beta + \beta \cdot x_\beta = \alpha(x_\alpha) + (\beta - \alpha)x_\beta.$$

We then compute the payment for an agent in A in the auditing auction:

$$\begin{aligned} p_i^A(t_i \mid 0) &= \frac{t_i \cdot x_\alpha}{(1 - t_i)} - \int_\alpha^{t_i} \frac{x_\alpha}{(1 - v)^2} \, \mathrm{d}v - \int_\beta^\alpha \frac{x_\beta}{(1 - v)^2} \, \mathrm{d}v \\ &= -x_\alpha + \frac{x_\alpha}{1 - \alpha} - \frac{x_\beta}{1 - \alpha} + \frac{x_\beta}{1 - \beta} = \frac{\alpha \cdot x_\alpha}{1 - \alpha} - \frac{x_\beta}{1 - \alpha} + \frac{x_\beta}{1 - \beta}. \end{aligned}$$

We can now compute the difference between these two payments as

$$\begin{aligned} p_i(t_i) - (1 - t_i)p_i^A(t_i \mid 0) &= x_\alpha \cdot \frac{\alpha(t_i - \alpha)}{1 - \alpha} + (\beta - \alpha)x_\beta - (1 - t_i)\left[-\frac{x_\beta}{1 - \alpha} + \frac{x_\beta}{1 - \beta}\right] \\ &= x_\alpha \cdot \frac{\alpha(t_i - \alpha)}{1 - \alpha} + x_\beta \left[\frac{\alpha^2 - \beta^2 + \alpha\beta^2 - \beta\alpha^2 + \beta t_i - \alpha t_i}{(1 - \beta)(1 - \alpha)}\right] \\ &= x_\alpha \cdot \frac{\alpha(t_i - \alpha)}{1 - \alpha} + x_\beta \left[\frac{-\alpha(t_i - \alpha)(1 - \beta) + \beta(t_i - \beta)(1 - \alpha)}{(1 - \beta)(1 - \alpha)}\right] \\ &= (x_\alpha - x_\beta)\frac{\alpha(t_i - \alpha)}{(1 - \alpha)} + x_\beta \frac{\beta(t_i - \beta)}{(1 - \beta)}. \end{aligned}$$

Substituting back in the probabilities and summing over the agents in A completes the proof. □

C Proofs from Section 6

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

$$\liminf_{n\to\infty}\left(\sum_{j=1}^{n}\mathbf{E}\left[a^j\cdot V(t^j,x(\hat{\sigma}(\hat{H}^j,j,t^j)))\right]-\sum_{j=1}^{n}\mathbf{E}\left[a^j\cdot V(t^j,x(\sigma(t^j)))\right]\right)\leq 0,$$

where $\hat{H}^j$ is history up to round j if the agent plays strategy $\hat{\sigma}$.

Proof. We first show the case when the allocation length ℓ is 1. We begin by showing that, for any $k > 1$, there exists an optimal strategy whose mapping at round k is independent of its history. First observe we need only consider histories in which $a^k = 1$ since if $a^k = 0$ the agent does not play any strategy. Consider any two such histories, denoted H^k and $(H')^k$. Assume the true history up until round k is $(H^1, ..., H^k)$ and call an optimal strategy σ_H. Consider also the strategy beginning at round k which plays as if the history preceding round k is $(H')^k$ and that this swap results in a strategy which is not optimal. We will refer to this strategy as $\sigma_{H'}$. We write the utility of each of these strategies as,

$$\sum_{j=1}^{N}\mathbf{E}\left[a^j\cdot V^j(\sigma_H^j)\right]=\sum_{j=1}^{k-1}a^j\cdot V(t^j,x(\sigma(H^j,j,t^j)))+\sum_{j=k}^{N}\mathbf{E}_{|a_k=1}\left[a^j\cdot V(t^j,x(\sigma(H^j,j,t^j)))\right]$$
$$\sum_{j=1}^{N}\mathbf{E}\left[a^j\cdot V^j(\sigma_{H'}^j)\right]=\sum_{j=1}^{k-1}a^j\cdot V(t^j,x(\sigma(H^j,j,t^j)))+\sum_{j=k}^{N}\mathbf{E}_{|a^k=1}\left[a^j\cdot V(t^j,x(\sigma(H^{j\prime},j,t^j))\right].$$

We can now compare the utility of these two strategies using our overtaking criterion and the optimality of $\sum_{j=1}^{N}V^j(\sigma_H^j)$. We must have

$$\sum_{j=1}^{k-1}a^j\cdot V(t^j,x(\sigma(H^j,j,t^j)))+\sum_{j=k}^{N}\mathbf{E}_{|a^k=1}\left[a^j\cdot V(t^j,x(\sigma(H^j,j,t^j)))\right]$$
$$\geq\sum_{j=1}^{k-1}a^jV(t^j,x(\sigma(H^j,j,t^j)))+\sum_{j=k}^{N}\mathbf{E}_{|a^k=1}\left[a^jV(t^j,x(\sigma(H^{j\prime},j,t^j)))\right].$$

The amount of value obtained prior to round k is identical for these two strategies so this reduces to

$$\sum_{j=k}^{N}\mathbf{E}_{|a^k=1}\left[a^jV(t^j,x(\sigma(H^j,j,t^j)))\right]\geq\sum_{j=k}^{N}\mathbf{E}_{|a^k=1}\left[a^jV(t^j,x(\sigma(H^{j\prime},j,t^j)))\right].$$

Lets now assume that the true history was $H^{k\prime}$ and we play the strategy corresponding to H^k. We can similarly derive

$$\sum_{j=k}^{N}\mathbf{E}_{|a^k=1}\left[a^jV(t^j,x(\sigma(H^{j\prime},j,t^j)))\right]\geq\sum_{j=k}^{N}\mathbf{E}_{|a^k=1}\left[a^jV(t^j,x(\sigma(H^j,j,t^j)))\right].$$

It is now obvious that in either of these cases both of these strategies provide the same utility. This contradicts the fact that swapping Histories results in a strategy which is not optimal. This can be done with any history that results in a different strategy being played. Therefore, no utility is lost by playing the same strategy at a given time regardless of the history.

We have shown that the optimal strategy is independent of the realized history at any round. We will now show that the optimal strategy is also independent of the round number. We will use a similar proof to the one used to prove history independence. Consider two rounds k and k' such that $\mathbf{E}[a^{k'}] > 0$ and $\mathbf{E}[a^k] > 0$. Let $\sigma_* = (\sigma_*^1, \ldots, \sigma_*^k, \ldots)$ be the optimal strategy. We define $\sigma_*^{\geq k}$ to be the portion of the optimal strategy that occurs after round k. We define $\sigma_*^{\geq k'}$ the same way. Now consider the strategy that plays $(\sigma_*^1, \ldots, \sigma_*^{\geq k})$ if $a^k = 0$ and $(\sigma_*^1, \ldots, \sigma_*^{\geq k'})$ if $a^k = 1$. We will call this strategy $\sigma_{k'}$. We can now compute the expected utility of these two strategies.

$$
\begin{aligned}
\sum_{j=1}^{N} \mathbf{E}\left[a^j \cdot V^j(\sigma_*^j)\right] =& \sum_{j=1}^{k-1} a^j V(t^j, x(\sigma_*(H^j, j, t^j))) \\
&+ \mathbf{Pr}\left[a^k = 1\right] \sum_{j=k}^{N} \mathbf{E}_{|a^k=1}\left[a^j \cdot V(t^j, x(\sigma_*(H^j, j, t^j)))\right] \\
&+ \mathbf{Pr}\left[a^k = 0\right] \sum_{j=k+1}^{N} \mathbf{E}_{|a^k=0}\left[a^j V(t^j, x(\sigma_*(H^j, j, t^j)))\right] \\
\sum_{j=1}^{N} \mathbf{E}\left[a^j \cdot V^j(\sigma_{k'}^j)\right] =& \sum_{j=1}^{k-1} a^j V(t^j, x(\sigma_*(H^j, j, t^j))) \\
&+ \mathbf{Pr}\left[a^k = 1\right] \sum_{j=k'}^{N+k'-k} \mathbf{E}_{|a^{k'}=1}\left[a^j V(t^j, x(\sigma_*(H^j, j, t^j)))\right] \\
&+ \mathbf{Pr}\left[a^k = 0\right] \sum_{j=k+1}^{N} \mathbf{E}_{|a^k=0}\left[a^j V(t^j, x(\sigma_*(H^j, j, t^j)))\right]
\end{aligned}
$$

Clearly these two strategies obtain the same amount of value prior to round k and the same amount of value if round k is a punishment round. We can again compare difference in these two utilities by utilizing the optimality of σ_*.

$$
\sum_{j=k}^{N} \mathbf{E}_{|a^k=1}\left[a^j \cdot V(t^j, x(\sigma(H^j, j, t^j)))\right] \geq \sum_{j=k'}^{N+k'-k} \mathbf{E}_{|a^k=1}\left[a^j \cdot V(t^i, x(\sigma(H^j, j, t^j)))\right]
$$

The right hand side of this inequality is exactly the utility an agent receives for playing $\sigma_*^{\geq k'}$ starting at round k' if round k' is an allocation round. It is now trivial to see that no utility is lost if at starting at time k' we play $\sigma_*^{\geq k}$ when k' is an allocation round. We can therefore replace $\sigma_*^{\geq k'}$ with $\sigma_*^{\geq k}$ when k' is an allocation round. Now if $\mathbf{E}[a^{k'} = 1] \neq 1$ we must also consider the case where k' is a punishment round. However we can just consider the strategy that plays $(\sigma_*^1, \ldots, \sigma_*^{\geq k})$ if $a^k = 1$ and $(\sigma_*^1, \ldots, \sigma_*^{\geq k'})$ if $a^k = 0$ and the rest of the proof follows the same steps as when k' is an allocation round.

□

Lemma 6.4. *Let $\{X_n, n \geq 1\}$ be an i.i.d. sequence of random variables such that $\mathbf{E}[|X_1|] < \infty$, and $\{I_n, n \geq 1\}$ another sequence of random variables which are allowed to be dependent on $\{X_n\}$ and such that I_n takes positive integer values and weakly increases with n, and almost surely goes to infinity with n. Then $\bar{X}_n := \frac{1}{I_n} \sum_{i=1}^{I_n} X_i$ converges almost surely to $\mathbf{E}[X_1]$.*

Proof. By the strong law of large numbers, $\frac{1}{n}\sum_{i=1}^{n} X_i$ converges almost surely to $\mathbf{E}[X_1]$. Let A be this event. Let B be the event that I_n goes to infinity with n. Then $\mathbf{Pr}[\overline{A \cap B}] = 0$. For any $\omega \in A \cap B$, $\{\frac{1}{I_n(\omega)}\sum_{i=1}^{I_n(\omega)} X_i(\omega)\}_n$ is a subsequence of $\{\frac{1}{n}\sum_{i=1}^{n} X_i(\omega)\}_n$, and therefore converges to $\mathbf{E}[X_1]$ as well. This shows that $\frac{1}{I_n}\sum_{i=1}^{I_n} X_i$ converges to $\mathbf{E}[X_1]$ almost surely. $\square$

Proposition 6.8. *A mechanism $\mathcal{M}$ satisfies BIC constraints if* (4) *holds for all mappings σ.*

Proof. The forward direction can be seen trivially by taking the expectation over the BIC constraints.

To see the reverse direction, assume that the mechanism satisfies the ex-ante truthfulness constraint but there exists two types $t^*, t' \in T$ that violate BIC constraints and therefore

$$t^* \cdot x(t') - p(t') > t^* \cdot x(t^*) - p(t^*).$$

Consider the mapping σ such that for all $t \neq t^*$, $\sigma(t) = t$ and $\sigma(t^*) = t'$. We must have that

$$\begin{aligned}
\mathbf{E}_t\left[t \cdot x(\sigma(t)) - p(\sigma(t)\right] &\leq \mathbf{E}\left[t \cdot x(t) - p(t)\right] \\
t^* \cdot x(\sigma(t^*)) - p(\sigma(t^*)) &\leq t^* \cdot x(t^*) - p(t^*) \\
t^* \cdot x(t') - p(t') &\leq t^* \cdot x(t^*) - p(t^*).
\end{aligned}$$

However this contradicts the assumption that BIC is violated on type t^*. $\square$

D Monotonicity of Virtual Utility

We must show that the distribution described by the density function $g(t_i) = \frac{e^{-t_i}}{1-e^{-1}}$ and cumulative density function $G(t_i) = \frac{1-e^{-t_i}}{1-e^{-1}}$ has audited virtual utility that strictly increases in t_i. In other words we must show that

$$\frac{(1 - G(t_i))(1 - \mathbf{E}_{b\sim G}[b \mid b \geq t_i])}{g_i(t_i)(1 - t_i)^2}$$

strictly increases in t_i. The virtual utility can also be written as

$$\frac{1 - G(t_i)}{g_i(t_i)} \cdot \frac{(1 - \mathbf{E}_{b\sim G}[b \mid b \geq t_i])}{(1 - t_i)^2}.$$

The first factor is

$$\frac{e^{-t_i} - e^{-1}}{e^{-t_i}} = 1 - e^{t_i - 1}.$$

We then calculate

$$\begin{aligned}
1 - \mathbf{E}_{b\sim G}\left[b \mid b \geq t_i\right] &= 1 - \frac{1}{1 - G(t_i)} \int_{t_i}^{1} g(v) v \, \mathrm{d}v \\
&= 1 - \frac{1 - e^{-1}}{e^{-t_i} - e^{-1}} \cdot \frac{e^{1-t_i}(1 + t_i) - 2}{e - 1} = \frac{1 - t_i e^{1-t_i}}{e^{1-t_i} - 1}.
\end{aligned}$$

Putting everything together, the audited virtual utility for type t_i is

$$\left(1 - e^{t_i - 1}\right) \cdot \frac{1 - t_i e^{1-t_i}}{e^{1-t_i} - 1} \cdot \frac{1}{(1 - t_i)^2} = \frac{e^{t_i - 1} - t_i}{(1 - t_i)^2}.$$

We plot this function for $t_i \in (0, 1)$ in Figure 1; it is indeed strictly increasing.

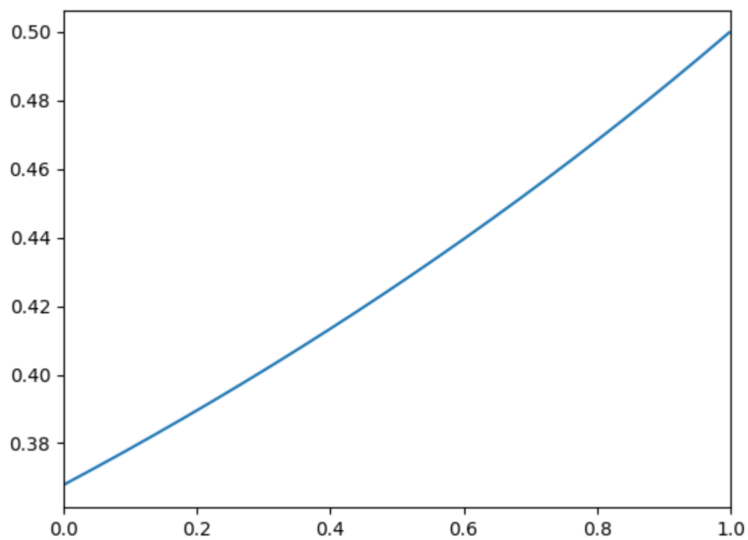

Figure 1: Audited virtual utility for the truncated exponential distribution.