# Prior-Free Dynamic Allocation Under Limited Liability

Sylvain Chassang Samuel Kapon[*][†]

Princeton University New York University

July 19, 2021.

**Abstract**

A principal seeks to efficiently allocate a productive public resource to a number of possible users. Vickrey-Clarke-Groves (VCG) mechanisms provide a detail-free way to do so provided users have deep pockets. In practice however, users may have limited resources. We study a dynamic allocation problem in which participants have limited liability: transfers are made ex post, and only if the productive efforts of participants are successful. We show that it is possible to approximate the performance of the pivot VCG mechanism using limited liability detail-free mechanisms that selectively ignore reports from participants who cannot make their promised payments. A complementary use of cautiousness and forgiveness achieves approximate renegotiation-proofness. We emphasize the use of prior-free online optimization techniques to approximate aggregate incentive properties of the pivot mechanism.

KEYWORDS: dynamic allocation, renegotiation proofness, lending, limited liability, VCG, pivot, approachability, online optimization, cautiousness, forgiveness.

[*]Chassang: chassang@princeton.edu, Kapon: sdk301@nyu.edu.

[†]We are grateful to the editor and three anonymous referees for their help in crafting the paper. We thank Gabriel Carroll, Rohit Lamba and Karl Schlag for helpful conversations. We are grateful to seminar audiences at Cornell, Microsoft Research, NYU and UCLA for their feedback.

# 1 Introduction

A principal repeatedly allocates a publicly managed resource, such as public land, radio spectrum, water, public space, low-interest loans, tax credits, or cash subsidies to several agents who can make productive use of this resource. When agents have quasi-linear preferences and private values, Vickrey-Clarke-Groves mechanisms (henceforth VCG, Vickrey (1961), Clarke (1971), Groves (1973)) can be used to allocate resources efficiently. An attractive property of VCG mechanisms is that they do not depend on the distribution of players' preferences. As such they are often referred to as prior-free or detail-free. However, VCG mechanisms also require agents to make upfront payments, which may not be feasible if agents do not have deep pockets. We study dynamic allocation under limited liability, i.e. when agents' payments are constrained by the stochastic output of their productive efforts.

Our benchmark model considers long-lived patient agents with quasilinear private preferences over an arbitrary set of policy choices. We are interested in mechanisms preserving the prior-free nature of VCG and allow the agents' values to follow an arbitrary exogenous stochastic process. We do not assume that the agents' values are i.i.d. or ergodic, precluding the possibility of learning the distribution of agents' preferences. Transfers are feasible at the end of each period, but limited to the resources produced by each agent during the period. This limited liability requirement makes our analysis relevant to environments where participants have limited access to credit. Examples range from the allocation of public resources such as water, electricity or land in developing communities, to the allocation of development funds, as well as credit lines to cities, states, or sovereign governments. In these examples, the agent receiving a public resource for productive use can only deliver on promised payments if the project for which the resource is used is successful.

The paper's main result describes prior-free limited liability allocation mechanisms that approximate the performance of the pivot VCG mechanism as the agents' horizon grows large. The allocation in the mechanism we construct selects a random subset of agents each period, a decision-group, and then chooses the efficient allocation of resources for this group. The key observation is that by selectively excluding from this group those agents who fail to make their externality payments, it is possible to: (i) keep a tight relationship between the players' aggregate transfers and their aggregate externality on others (*incentive alignment*); (ii) allocate resources efficiently with large probability (*efficient allocation*). This is achieved by treating the incentive approximation problem as a regret minimization problem which can be solved in a prior-free manner using online optimization methods (Blackwell, 1956, Hannan, 1957, Foster and Vohra, 1999, Cesa-Bianchi and Lugosi, 2006). As the horizon becomes large, the resulting mechanism implements efficient allocation in $\epsilon$-Nash equilibrium.

An important practical concern is the credibility of commitment: if the mechanism results in large losses in efficiency after rare histories, perhaps it is implausible that the principal will stick with the mechanism at such histories. We show how to address the issue of commitment by expanding the set of target incentive properties. By requiring *incentive alignment* and *efficient allocation* to hold from the perspective of any history, it is possible to implement approximately efficient allocation in perfect $\epsilon$-equilibrium (Radner, 1980). In a model with discounting, a similar mechanism implements efficient allocation in contemporaneous perfect $\epsilon$-equilibrium in the sense of Mailath et al. (2005). This ensures that continuing allocations are approximately efficient after any history, so that the principal and agent have limited incentives to renegotiate the rules of the mechanism over time. The corresponding mechanism exhibits a mix of cautiousness and forgiveness: as agents start missing their externality payments, they are excluded from the decision group with increasing probability, thereby limiting the extent of efficiency loss they can cause; in turn, even if agents make consistently erroneous claims over a long interval of time, they are still included in the decision group with positive probability and they get swiftly re-included in the decision-making group if they resume making their externality payments. In other words, the mechanism does not rely on grim trigger punishments, and does not lead to efficiency traps, even after lengthy deviations from truthful reporting.

The paper lies at the intersection of different strands in the literature on dynamic mechanism design. Work by Casella (2005), Jackson and Sonnenschein (2007) and Escobar and Toikka (2009) has shown how to build mechanisms achieving efficient dynamic allocation in environments without transferability.[1] However, these mechanisms rely crucially on the assumption that the state of the world is ergodic. Under truthful behavior, the sample distribution of messages over time must match the prior distribution of states. One can implement approximately efficient allocations by constraining the realized joint distribution of agents' messages to match the prior joint distribution of agent's messages, and selectively excluding messages when the sample distribution departs from the anticipated distribution. This paper assumes partial transferability, up to a limited liability constraint, but relaxes the assumption that the state of the world follows an ergodic process. The realized distribution of states may differ from the expected distribution of states with large probability.

In a related line of inquiry, Bergemann and Välimäki (2010) and Athey and Segal (2013) study dynamic allocation problems with fully transferable payoffs. Crucially, they allow future states to depend on past allocations. As a result, a player's externality must incorporate expected impacts on future periods, and the corresponding mechanism is not detail-free. The current paper imposes limited liability constraints and considers detail-free mechanisms, but assumes that the process for participants' values is exogenous.

Finally, this paper shares both the methods and concerns of Chassang (2013) which shows how to dynamically approximate a high-liability single-agent incentive contract under limited liability constraints. The current paper differs by considering a multi-agent allocation problem, rather than a single-agent incentive problem. As a result, the target incentive properties as well as the corresponding approximation strategy are significantly different. We also address two important concerns absent from Chassang (2013): we extend the analysis

[1] Olszewski and Safronov (2018a,b) exhibit explicit repeated game equilibria implementing efficient allocations, replicating externality payments using continuation values. Athey and Miller (2007) also studies efficient dynamic allocation in an infinitely repeated setting, and also emphasizes limits on the agents' ability to make transfers.

to infinite horizon games with discounting, and we deal with environments in which we relax the assumption that an agent observes outcomes for all actions including those not taken by the principal, which we refer to as counterfactual outcomes. Although the infinite horizon setting has advantages, the simplest application of the techniques from the online optimization literature is to the undiscounted finite horizon setting, and the core ideas are clearest in this case. For this reason, we develop our baseline mechanism in the undiscounted finite horizon setting, and later show how it extends to the discounted infinite horizon case.

The paper is structured as follows. Section 2 introduces the framework. Section 3 describes the usual pivot mechanism in our repeated context and discusses equilibrium multiplicity issues. Section 4 introduces our benchmark mechanism, and shows how to approximate incentive properties using methods from online optimization. Section 5 establishes that our baseline mechanism extends essentially as is when output is not observed by the principal. Section 6 shows how to address renegotiation-proofness. Section 7 extends the analysis to infinite horizon games with discounting. Appendix A deals with unobserved counterfactuals, while Appendix B collects proofs omitted from the main text.

## 2 Framework

**Decisions, transfers, and payoffs.** In each period $t \in \{1, \cdots, N\}$ a principal picks a decision (sometimes referred to as allocation) $a_t \in A$ affecting the productivity of a finite number of agents indexed by $i \in I$. In any period $t$, a decision $a \in A$ induces stochastic outputs $(y_{i,t}(a))_{i \in I}$ for each player.[2] By assumption, $y_{i,t}(a) \in [0, y_{\max}]$, with $y_{\max}$ a fixed upper bound. We denote by $y_{i,t} \equiv (y_{i,t}(a))_{a \in A}$ the tuple of possible outputs for agent $i$ in period $t$ for different decisions $a \in A$ taken by the principal. Let $y_t \equiv (y_{i,t})_{i \in I}$ be the profile of possible outputs across players.

Agents are able to make transfers $\tau_{i,t}$ to the principal, but are limited by their output:

[2] We abuse notation and denote $y_{i,t}(a)$ for both random variable and its realized value.

$\tau_{i,t} \in [0, y_{i,t}(a_t)]$. This restriction can be thought of as a credit constraint, along the lines of Che and Gale (1998), except that the constraint is linked to the realized output of the agent. The relevant constraint here is that transfers are bounded above by realized returns: if an agent claims a public resource but fails to generate any output, they are unable to make transfers in this period.[3]

Players are patient and do not discount time, so that player $i$'s aggregate utility boils down to

$$\sum_{t=1}^{N} y_{i,t}(a_t) - \tau_{i,t}.^{4} \tag{1}$$

The principal is risk neutral and seeks to maximize total output $\sum_{t=1}^{N} \sum_{i \in I} y_{i,t}(a_t)$.

**Agents' Information.** We consider private value environments. At the beginning of each period $t$, each agent $i$ observes their private value $v_{i,t}$ over decisions $a \in A$: $v_{i,t}(a) = \mathbb{E}[y_{i,t}(a)|\mathcal{F}_{i,t}]$, where $\mathcal{F}_{i,t}$ represent the information of player $i$ at the beginning of period $t$.

We assume throughout the main text that at the end of each period $t$, each agent observes the output $y_{i,t}(a)$ for all possible decisions of the principal, $a \in A$. Appendix A extends the analysis to cases in which agent $i$ observes only the output $y_{i,t}(a_t)$ corresponding to the actual decision $a_t$.

**Assumption 1 (private values)** *Agents' values are sufficient statistic for their output at time $t$: for all $i$, $t$,*

$$\mathbb{E}[y_{i,t}|(\mathcal{F}_{j,t})_{j \in I}] = \mathbb{E}[y_{i,t}|\mathcal{F}_{i,t}] = v_{i,t}.$$

*The stochastic process for private values $(v_{i,t})_{i \in I, t \geq 1}$ is exogenously given, and does not depend on past allocation decisions.*

[3]Savings would relax this constraint. See Section 8 for a discussion.
[4]We extend the analysis to an infinite-horizon version of the model with discounting in Section 7.

An implication of Assumption 1 is that the allocation maximizing expected output for any subgroup $G \subset I$ of agents is a function of their current values alone.

**Principal's Information.** The principal does not observe $(v_{i,t})_{i \in I}$. For clarity, we develop our analysis in steps. In Section 4, we assume that the principal observes $y_t$. We relax this assumption in Section 5 and show that the principal can instead rely on reported output.

**Mechanisms.** In each period $t$, agent $i$ sends a messages $m_{i,t} \in M_i$ to the principal. We denote by $m_t \equiv (m_{i,t})_{i \in I}$ the profile of messages. A mechanism maps the history of messages and observed outputs to a stochastic process $(a_t, (\tau_{i,t})_{i \in I})_{t \in \{1, \cdots, N\}}$ of allocations and transfers adapted to the information available to the principal.

**Solution concepts.** Any mechanism induces a game $\Gamma$ associating reporting processes $m_i = (m_{i,t})_{t \geq 1}$ for each agent $i$ to payoffs

$$\gamma(m_i, m_{-i}) = \frac{1}{N} \mathbb{E} \left[ \sum_{t=1}^{N} y_{i,t}(a_t) - \tau_{i,t} \right].$$

We consider three main solution concepts: $\epsilon$-Nash equilibrium, perfect $\epsilon$-Nash equilibrium (Radner, 1980), and following the critique of Mailath et al. (2005), contemporaneous perfect $\epsilon$-equilibrium.

# 3 Benchmark Results

In this section, we briefly relax the limited liability constraint so that VCG payments are feasible. We describe the usual pivot VCG mechanism in our context, and highlight issues of equilibrium multiplicity that arise in repeated settings. The same issues apply to the limited liability mechanisms we study in later sections. We then briefly discuss the challenges presented by limited liability.[5]

[5] We discuss difficulties in extending our analysis to VCG mechanisms other than the pivot in Section 8.

Since there are no intertemporal externalities, the dynamic pivot mechanism is no different from the static pivot mechanism. It consists of requesting messages $m_{i,t}$ corresponding to players' values: $m_{i,t} \in V_i = [0, y_{\max}]^A$. For any group of agents $G \subset I$ and message profile $m_t$, we define

$$a^*(m_t|G) \in \arg\max_{a \in A} \sum_{i \in G} m_{i,t}(a)$$

the efficient allocation for group $G$ given reports $m_t$.

The allocation is set to $a^*(m_t|I)$: the efficient allocation for the overall group $I$ given stated values $m_t$. Let $I \setminus i$ denote the set of agents other than $i$. Transfers from agent $i$ are set to

$$\tau_{i,t} = \sum_{j \neq i} m_{j,t}(a^*(m_t|I \setminus i)) - m_{j,t}(a^*(m_t|I)),$$

i.e., agent $i$'s reported externality on others.

**Proposition 1** *For all $i, t$, transfers are positive: $\tau_{i,t} \geq 0$. Truthful revelation, i.e. $m_{i,t} = v_{i,t}$ for all $i, t$, is a perfect Bayesian equilibrium which implements efficient allocation.*

**Equilibrium multiplicity.** As in the case of the static pivot mechanism, there may exist other equilibria.[6] In particular, the dynamic nature of the game, combined with the fact that players are transferring a significant amount of surplus to the principal, creates scope for collusive strategies among bidders.[7] For instance, along the lines of many real cartels, bidders may use bid rotation supported by a reversion to stage game Nash (Aoyagi, 2003, Skrzypacz and Hopenhayn, 2004). This would result in inefficient allocation.

While it is possible to leverage the finite horizon assumption to ensure that collusive strategies unravel, we believe that this observation provides limited solace. When horizon

[6]For instance, in auctions, bidders who know they will be losing in equilibrium may misrepresent their values.

[7]The fact that surplus extraction in repeated mechanisms creates scope for collusion has received some attention. Abdulkadiroglu and Chung (2003) solve for the optimal auction design under the assumption that players choose the equilibrium that minimizes revenues for the auctioneer. Lee and Sabourian (2011) provide conditions for full implementation of efficient allocations in an i.i.d. setting.

$N$ is large, collusive strategies remain a perfect $\epsilon$-equilibrium in the undiscounted game (and would be full fledged perfect Bayesian equilibria in the infinite horizon game with discounting). Reputational arguments à la Kreps et al. (1982) also suggest that collusion may be plausible.

This observation helps clarify why we are only concerned in implementing efficient allocation in some equilibrium, rather than all equilibria. This is an important limitation compared to Jackson and Sonnenschein (2007) who are able to establish unilateral payoff-guarantees corresponding to those under efficient allocation.

**Difficulties introduced by limited liability.** Under the pivot mechanism, agents get positive surplus in expectation: $\mathbb{E}[y_{i,t}(a_t) - \tau_{i,t}] \geq 0$. If there was no uncertainty about the realization of returns conditional on allocation, this would imply that in equilibrium $y_{i,t}(a_t) \geq \tau_{i,t}$. In other terms, the constraint that transfers are bounded above by realized returns would be non-binding. In contrast, when returns are uncertain, the pivot mechanism may require the agent to make a positive transfer $\tau_{i,t} > 0$ even though the agent's realized output $y_{i,t}(a_t)$ is, for instance, equal to zero. Note that the timing is important: allocation must happen before returns are realized. As a result, the limited liability constraint, $\tau_{i,t} \leq y_{i,t}(a_t)$, can bind whenever returns are uncertain conditional on allocation.

An intuitive fix would be to let agents accrue debt if they cannot make their externality payment, and reimburse past debts whenever they can. This is indeed a component of the mechanisms we construct. However, this fix allows an agent to repeatedly influence the allocation, even if they repeatedly fail to make their externality payments. As a result, this simple mechanism does not generalize well, for instance to the case where the principal does not observe the output. Rather, agents must be punished when they repeatedly fail to make their externality payments, otherwise they may profit by overclaiming value for their preferred allocation and subsequently underclaiming realized output. The difficulty is that punishment may come at the cost of efficiency losses. The value of the mechanisms studied in the remainder of the paper is to deliver adequate incentives without sacrificing significant allocative efficiency.

# 4 A Limited Liability Mechanism

We now construct a limited-liability prior-free mechanism that implements efficient allocation in $\epsilon$-Nash equilibrium. The construction proceeds in three step: first, we describe target incentive properties that our mechanism should satisfy; second, we show how we can satisfy these properties by viewing them as a regret minimization problem; third, we show that under the resulting mechanism, truthful reporting of one's value is an $\epsilon$-Nash equilibrium inducing efficient allocation.

## 4.1 An explicit mechanism

Each period, agent $i$'s message space is the set of possible flow values $v_i$ over allocations $a$: $M_i = V_i = [0, y_{\max}]^A$. We assume in this section that $y_t$ is observable to the principal. We denote by $h_t = (m_1, y_1, \cdots, y_{t-1}, m_t)$ the history of messages and outputs up to the decision stage of period $t$. Our mechanism is specified by the following objects:

- Each period, a decision group $I_t \subset I$ is picked according to a distribution $\mu_t \in \Delta(\mathcal{P}(I))$, where $\mathcal{P}(I)$ is the set of subsets of $I$. Distribution $\mu_t$ depends only on history $h_t$. The principal implements allocation $a_t = a^*(m_t|I_t)$, i.e. the optimal allocation for group $I_t \subset I$ given reported preferences $m_t$.
- A feasible set of transfers $(\tau_{i,t}(a))_{i\in I}$ is implemented as a function of the allocation $a$ chosen by the principal.

**Target properties.** We seek to specify processes $\mu_t$, and $\tau_{i,t}$ so that they replicate key properties of the pivot mechanism: allocations should be approximately efficient, and transfers $\tau_{i,t}$ should ensure that players internalize their externality on others.

It is helpful to introduce the following notation. For any group $G \subset I$, and distribution $\mu \in \Delta(\mathcal{P}(I))$ over the set $\mathcal{P}(I)$ of subsets of $I$, we introduce the following notation.

- Aggregate outputs, when group $G$ is the decision group, and when the decision group is picked according to $\mu$:

$$Y_t(G) \equiv \sum_{i \in I} y_{i,t}(a^*(m_t|G)) \quad \text{and} \quad Y_t(\mu) \equiv \sum_{G \in \mathcal{P}(I)} \mu(G) Y_t(G). \tag{2}$$

- Output for all agents excluding $i$, when group $G$ is the decision group, and when the decision group is picked according to $\mu$:

$$Y_{-i,t}(G) \equiv \sum_{j \in I \setminus i} y_{j,t}(a^*(m_t|G)) \quad \text{and} \quad Y_{-i,t}(\mu) \equiv \sum_{G \in \mathcal{P}(I)} \mu(G) Y_{-i,t}(G). \tag{3}$$

- Expected transfer, and expected output for agent $i$ when the decision group is picked according to $\mu$:

$$\tau_{i,t}(\mu) \equiv \sum_{G \in \mathcal{P}(I)} \mu(G) \tau_{i,t}(a^*(m_t|G)) \quad \text{and} \quad y_{i,t}(\mu) \equiv \sum_{G \in \mathcal{P}(I)} \mu(G) y_{i,t}(a^*(m_t|G)).$$

In particular $Y_t(I)$, $Y_{-i,t}(I)$ and $Y_{-i,t}(I \setminus i)$ respectively denote total output for group $I$ when picking the reportedly optimal allocation for group $I$, total output for group $I \setminus i$ when picking the reportedly optimal allocation for group $I$, and total output for group $I \setminus i$ when picking the reportedly optimal allocation for group $I \setminus i$.

We seek to ensure the following properties, where the usual "little o" notation $o(T)$ denotes terms "negligible compared to $T$ as $T$ becomes large."

$$\text{(efficient allocation)} \qquad \mathcal{R}_T^I \equiv \sum_{t=1}^{T} Y_t(I) - Y_t(\mu_t) \leq o(T); \tag{4}$$

$$\text{(incentive alignment)} \quad \forall i \in I, \quad \mathcal{R}_{i,T}^{\tau} \equiv \sum_{t=1}^{T} Y_{-i,t}(I \setminus i) - Y_{-i,t}(\mu_t) - \tau_{i,t}(\mu_t) \leq o(T). \tag{5}$$

Regrets $\mathcal{R}_T^I$ and $\mathcal{R}_{i,T}^{\tau}$ respectively measure the extent to which the allocation generated via $(\mu_t)_{t \in \mathbb{N}}$ is efficient and transfers $(\tau_{i,t})_{t \in \mathbb{N}}$ cover agent $i$'s externality on other participants.

**Mechanism design via regret minimization.** We use tools from the online optimization literature (Blackwell, 1956, Hannan, 1957) to control regrets in a prior-free way. Define expected profits $\pi_{i,t}(\mu_t) = y_{i,t}(\mu_t) - \tau_{i,t}(\mu_t)$. Regret $\mathcal{R}^{\tau}_{i,T}$ can be written as

$$\mathcal{R}^{\tau}_{i,T} = \sum_{t=1}^{T} \pi_{i,t}(\mu_t) - y_{i,t}(I \setminus i) + \sum_{t=1}^{T} Y_t(I \setminus i) - Y_t(\mu_t).$$

Consider the vector of regrets $\mathcal{R}_T \equiv (\mathcal{R}^I_T, \mathcal{R}^{\tau}_{i,T})_{i \in I}$, including both aggregate efficiency regrets and incentive regrets for each player. Following Blackwell (1956) we choose transfers $\tau_{i,t}$ and $\mu_t$ so that the approachability condition

$$\left\langle \mathcal{R}^+_T, \mathcal{R}_{T+1} - \mathcal{R}_T \right\rangle \leq 0 \tag{6}$$

is satisfied for all realized values of outputs $(y_{i,T+1})_{i \in I}$. The logic behind condition (6) is that it ensures that marginal regrets $\mathcal{R}_{T+1} - \mathcal{R}_T$ go in a direction opposite, or at least orthogonal, to the positive part of accumulated regrets $\mathcal{R}^+_T$, thereby ensuring that regrets do not accumulate too fast. We have that:

$$\begin{aligned} \left\langle \mathcal{R}^+_T, \mathcal{R}_{T+1} - \mathcal{R}_T \right\rangle = {} & \overbrace{\sum_{i \in I} \left[\mathcal{R}^{\tau}_{i,T}\right]^+ \left(\pi_{i,T+1}(\mu_t) - y_{i,T+1}(I \setminus i)\right)}^{A} \\ & + \underbrace{\left[\mathcal{R}^I_T\right]^+ \left(Y_{T+1}(I) - Y_{T+1}(\mu_{T+1})\right) + \sum_{i \in I} \left[\mathcal{R}^{\tau}_{i,T}\right]^+ \left(Y_{T+1}(I \setminus i) - Y_{T+1}(\mu_{T+1})\right)}_{B} \end{aligned} \tag{7}$$

Set distribution $\mu_{T+1}$ so that for all $i \in I$,

$$\mu_{T+1}(I \setminus i) = \frac{\left[\mathcal{R}^{\tau}_{i,T}\right]^+}{\left[\mathcal{R}^I_T\right]^+ + \sum_{i \in I} \left[\mathcal{R}^{\tau}_{i,T}\right]^+} \tag{8}$$

and $\mu_{T+1}(I) = 1 - \sum_{i \in I} \mu_{T+1}(I \setminus i)$.[8] This ensures that term $B$ of (7) is equal to 0.

[8] If all regrets are negative, then $\mu_{T+1}(I) = 1$.

In turn, set transfers $\tau_{i,T+1}(a)$ given decision $a \in A$ so that

$$\tau_{i,T+1}(a) = \begin{cases} y_{i,T+1}(a) & \text{if} \quad \mathcal{R}^{\tau}_{i,T} > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

This ensures that the term $A$ of (7) is less than or equal to 0.

The transfer scheme described by condition (9) corresponds to the intuitive fix of the standard pivot mechanism: let agents accumulate debts, and have them repay their debt as much as possible whenever possible.[9] The distribution $\mu$ over decision-making groups defined by (8) balances the need to provide agents with correct incentives, and the desire to maintain approximately efficient allocation. Agents who fail to make their externality payments are ignored provided the efficiency losses from doing so are not too large.

**Proposition 2** *Consider the mechanism $(\mu_t, (\tau_{i,t})_{i\in I})_{t\geq 1}$ defined by (8) and (9). Then, for all $\epsilon > 0$, there exists $N_0$ such that for all $N > N_0$, truthful revelation, i.e. $(m_{i,t})_{i\in I, t\geq 1} = (v_{i,t})_{i\in I, t\geq 1}$, is an $\epsilon$-Nash equilibrium of the induced game $\Gamma$. Under truthful reporting, the allocation approaches efficiency as the horizon $N$ gets large.*

**Proof:** We begin by showing that $||\mathcal{R}^+_N|| = o(N)$, where $||\cdot||$ denotes the Euclidean norm over vectors. Using the fact that process $(\mu_t, \tau_{i,t})_{i\in I, t\geq 1}$ satisfies approachability condition (6) for all $T$, we have that

$$\begin{aligned} ||\mathcal{R}^+_{T+1}||^2 &\leq ||\mathcal{R}^+_T||^2 + 2\left\langle \mathcal{R}^+_T, \mathcal{R}_{T+1} - \mathcal{R}_T \right\rangle + ||\mathcal{R}_{T+1} - \mathcal{R}_T||^2 \\ &\leq ||\mathcal{R}^+_T||^2 + ||\mathcal{R}_{T+1} - \mathcal{R}_T||^2 \leq ||\mathcal{R}^+_T||^2 + (|I|+1)^3 y^2_{\max} \\ &\leq (T+1)(|I|+1)^3 y^2_{\max}. \end{aligned}$$

[9] Note that in this specification, transfers $\tau_{i,t}$ correspond to the entire output whenever the agent has any debt (i.e. $\mathcal{R}^{\tau}_{i,t-1} > 0$). In some circumstances this payment may be greater than the agent's debt. Because payoffs are averaged over time, this does not matter. Excess payments in period $t$ imply $\mathcal{R}^{\tau}_{i,t} < 0$ so that the agent is allowed to skip some externality payments in the future. In an alternative specification, the transfer could be set to the debt owed by the agent. This does not affect our results.

It follows that $||\mathcal{R}_N^+|| = O(\sqrt{N})$. In addition, we now show that under truthful reporting by all agents, the assumption of private values implies that for all $i \in I$,

$$\mathbb{E}\left[\mathcal{R}_{i,N}^{\tau}\right] \geq -o(N). \tag{10}$$

Consider the negative part of payoff regrets $\mathcal{R}_{i,T}^{\tau,-} \equiv \max\{0, -\mathcal{R}_{i,T}^{\tau}\}$. We have that

$$\mathbb{E}\left[[\mathcal{R}_{i,T+1}^{\tau,-}]^2\right] \leq \mathbb{E}\left[[\mathcal{R}_{i,T}^{\tau,-}]^2\right] - 2\mathbb{E}\left[\left\langle \mathcal{R}_{i,T}^{\tau,-}, \mathcal{R}_{i,T+1}^{\tau} - \mathcal{R}_{i,T}^{\tau}\right\rangle\right] + \mathbb{E}\left[[\mathcal{R}_{i,T+1}^{\tau} - \mathcal{R}_{i,T}^{\tau}]^2\right]. \tag{11}$$

Note that $\mathcal{R}_{i,T+1}^{\tau} - \mathcal{R}_{i,T}^{\tau} = Y_{-i,T+1}(I\backslash i) - Y_{-i,T+1}(\mu_{T+1}) - \tau_{i,T+1}(\mu_{T+1})$, and that $\tau_{i,T+1} = 0$ whenever $\mathcal{R}_{i,T}^{\tau,-} > 0$. In addition, the assumption of private values implies that

$$\mathbb{E}\left[Y_{-i,T+1}(I\backslash i) - Y_{-i,T+1}(\mu_{T+1}) \,\middle|\, m_{-i,T+1} = v_{-i,T+1}\right] > 0.$$

This implies that $-2\mathbb{E}\left[\left\langle \mathcal{R}_{i,T}^{\tau,-}, \mathcal{R}_{i,T+1}^{\tau} - \mathcal{R}_{i,T}^{\tau}\right\rangle\right] \leq 0$, so that (11) yields

$$\begin{aligned}\mathbb{E}\left[[\mathcal{R}_{i,T+1}^{\tau,-}]^2\right] \leq& \mathbb{E}\left[[\mathcal{R}_{i,T}^{\tau,-}]^2\right] + \mathbb{E}\left[[\mathcal{R}_{i,T+1}^{\tau} - \mathcal{R}_{i,T}^{\tau}]^2\right] \\ \leq& \mathbb{E}\left[[\mathcal{R}_{i,T}^{\tau,-}]^2\right] + I^2 y_{\max}^2 \leq (T+1)|I|^2 y_{\max}^2.\end{aligned}$$

Jensen's inequality implies $\mathbb{E}\left[\mathcal{R}_{i,N}^{\tau,-}\right] \leq \sqrt{\mathbb{E}\left[\left[\mathcal{R}_{i,N}^{\tau,-}\right]^2\right]} \leq |I| y_{\max}\sqrt{N}$, which yields (10).

We now show that for $N$ sufficiently large, player $i$ can benefit at most by $\epsilon$ from deviating from truthful reporting. Observe that (4) and (5) hold with $o(T)$ dependent only on $T$, $|I|$ and $y_{\max}$, which implies that (4) and (5) hold in expectation with the same $o(T)$ term for any messaging strategies.

As a result, (5) implies that for any messaging strategy $m_i$,

$$\mathbb{E}_{m_i, v_{-i}}\left[\sum_{t=1}^{N} y_{i,t}(\mu_t) - \tau_{i,t}(\mu_t)\right] \leq \mathbb{E}_{m_i, v_{-i}}\left[\sum_{t=1}^{N} Y_{I,t}(\mu_t) - Y_{-i,t}(I\backslash i)\right] + o(N)$$

$$\leq \mathbb{E}_{v_i, v_{-i}} \left[ \sum_{t=1}^{N} Y_{I,t}(I) - Y_{-i,t}(I \backslash i) \right] + o(N).$$

where $o(N)$ is independent of $m_i$. In turn, using (10) and (4), it follows that under truthtelling player $i$ can achieve a payoff

$$\mathbb{E}_{v_i, v_{-i}} \left[ \sum_{t=1}^{N} y_{i,t}(\mu_t) - \tau_{i,t}(\mu_t) \right] \geq \mathbb{E}_{v_i, v_{-i}} \left[ \sum_{t=1}^{N} Y_{I,t}(\mu_t) - Y_{-i,t}(I \backslash i) \right] - o(N)$$
$$\geq \mathbb{E}_{v_i, v_{-i}} \left[ \sum_{t=1}^{N} Y_{I,t}(I) - Y_{-i,t}(I \backslash i) \right] - o(N).$$

The result that truthful revelation is an $\epsilon$-Nash equilibrium follows for $N$ large enough that $o(N) < \epsilon N$. Given truthful revelation, the fact that the allocation approaches efficiency follows from condition (4). ■

The mechanism described by (8) and (9) can be viewed as a lending protocol. It ensures that agents pay for their externality on others by selectively ignoring the preferences of participants who failed to make the necessary payments. The key observation allowing (5) and (4) to be satisfied together is that an agent who repeatedly fails to make adequate externality payments can also be ignored from the perspective of allocative efficiency.

One important aspect of our lending protocol is that the credit line available to players is modulated by the counterfactual performance gains from including them into the decision process, captured by efficiency regret $\mathcal{R}_T^I$. A participant who is excluded from the decision process one period may be reincluded if doing so retroactively would have led to performance gains.[10] In Section 6, we build on this feature to ensure that our mechanism is renegotiation-proof.

Before moving to renegotiation-proofness, we show in the next section that our mechanism extends essentially as is when the principal does not observe outputs and instead relies on the agents' reports of both values and output.

[10] Appendix A shows that this construction can be extended to settings with unobservable counterfactuals.

# 5 Strategic Feedback

We consider now the case where outputs $(y_{i,t}(a))_{a\in A}$ are not observed by the principal, but rather are privately observed by agent $i$. This maintains the assumption that agents observe their own counterfactual outcomes, even though the principal does not.[11] As a result, a message $m_{i,t}$ now consists of stated values $\overline{v}_{i,t}$ and stated outputs $\overline{y}_{i,t-1}$. We denote by $\overline{Y}_t$ and $\overline{Y}_{-i,t}$ the reported analogues of $Y_t$ and $Y_{-i,t}$ defined in (2) and (3).

**Target properties.** Target properties are identical to those defined in Section 4, replacing realized outputs with reported outputs. In each period, for each agent $i$ the mechanism chooses a distribution $\mu_t \in \Delta(\mathcal{P}(I))$ defining a random subset $I_t \subset I$ and transfers $\tau_{i,t} \in [0, \overline{y}_{i,t}(a_t)]$. The objective is to satisfy the following target properties.

$$\text{(incentive alignment)} \quad \forall i, T, \quad \overline{\mathcal{R}}^{\tau}_{i,T} = \sum_{t=1}^{T} \overline{Y}_{-i,t}(I\backslash i) - \overline{Y}_{-i,t}(\mu_t) - \tau_{i,t}(\mu_t) \leq o(T) \tag{12}$$

$$\text{(efficient allocation)} \quad \forall T, \quad \overline{\mathcal{R}}^{I}_{T} = \sum_{t=1}^{T} \overline{Y}_{t}(I) - \overline{Y}_{t}(\mu_t) \leq o(T) \tag{13}$$

Define $\overline{\pi}_{i,t}(\mu) \equiv \overline{y}_{i,t}(\mu) - \tau_{i,t}$. We have that

$$\overline{\mathcal{R}}^{\tau}_{i,T} = \sum_{t=1}^{T} \overline{\pi}_{i,t}(\mu_t) - \overline{y}_{i,t}(I\backslash i) + \overline{Y}_{t}(I\backslash i) - \overline{Y}_{t}(\mu_t).$$

Let $\overline{\mathcal{R}}_T \equiv \left(\overline{\mathcal{R}}^{I}_{T}, \overline{\mathcal{R}}^{\tau}_{i,T}\right)_{i\in I}$. Let $\mu_{T+1}$ be the distribution over $\{I, I\setminus i \mid i \in I\}$ such that

$$\forall i \in I, \mu_{T+1}(I\backslash i) = \frac{\left[\overline{\mathcal{R}}^{\tau}_{i,T}\right]^{+}}{\left[\overline{\mathcal{R}}^{I}_{T}\right]^{+} + \sum_{j\in I}\left[\overline{\mathcal{R}}^{\tau}_{j,T}\right]^{+}}.$$

[11]We show how to relax this assumption in Appendix A.

Transfers are defined by

$$\tau_{i,T+1}(a_{T+1}) = \begin{cases} \overline{y}_{i,T+1}(a_{T+1}) & \text{if} \quad \overline{\mathcal{R}}^{\tau}_{i,T} > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Altogether, this ensures that for all output realizations, the approachability condition holds:

$$\left\langle \overline{\mathcal{R}}^{+}_{T}, \overline{\mathcal{R}}_{T+1} - \overline{\mathcal{R}}_{T} \right\rangle \leq 0.$$

In turn this implies that conditions (12) and (13) hold.

**Proposition 3** *For all $\epsilon > 0$, there exists $N_0$ such that for all $N > N_0$, truthful revelation $m_{i,t} = (v_{i,t}, y_{i,t-1})$ is an $\epsilon$-Nash equilibrium of $\Gamma$.*

The proof is similar to the proof of Proposition 2 and is given in Appendix B.1. Incentive alignment ensures that an agent cannot benefit from underreporting realized output since this would lead to a commensurate reduction in the agent's impact on future decision-making.

## 6 Dynamic Consistency

The mechanism of Section 4 fails two important forms of dynamic consistency. First, it is not renegotiation proof. Imagine that a player keeps failing to make externality payments. This could be due to rare bad luck, intentional misreporting, or some systematic reporting error. In principle, the mechanism of Section 4 may exclude the player for a large number of periods going forward. This is inefficient, and the principal may be persuaded to reset records and start taking into account the player's preferences again.

Second, truthful revelation may no longer be an $\epsilon$-Nash equilibrium after histories that are rare but not impossible. For instance, a player $i$ may end up having a large positive externality on others: it is unlikely but possible that $\sum_{t=1}^{T} Y_{-i,t}(\mu_t) - Y_{-i,t}(I \setminus i) \gg 0$. After such histories, player $i$'s incentives for truthtelling are severly weakened.

The mechanism presented in this section achieves approximate dynamic consistency by constructing a measure of externalities that smoothes out large positive deviations, and by requiring that the efficient allocation requirement (4) hold starting from any period.

Drawing on Radner (1980), we formalize our criteria for dynamic consistency as follows. For any history $h_T$ and profile of reporting strategies $m = (m_i)_{i \in I}$, we define continuation payoffs to player $i$, and continuation surplus as

$$\gamma_i(m_i, m_{-i}|h_T) = \mathbb{E}\left[\frac{1}{N}\sum_{t=T}^{N} y_{i,t} - \tau_{i,t}\middle| h_T\right] \quad \text{and} \quad S(m|h_T) = \mathbb{E}\left[\frac{1}{N}\sum_{t=T}^{N}\sum_{i \in I} y_{i,t}\middle| h_T\right]$$

**Definition 1 (dynamic consistency)** *Pick $\epsilon > 0$. Strategy profile $m = (m_i)_{i \in I}$ is a perfect $\epsilon$-equilibrium if and only if for all histories $h_t$*

$$\forall \widehat{m}_i, \quad \gamma_i(m_i, m_{-i}|h_t) + \epsilon \geq \gamma_i(\widehat{m}_i, m_{-i}|h_t).$$

*Strategy profile $m = (m_i)_{i \in I}$ is $\epsilon$-renegotiation proof if and only if for all histories $h_t$ and all alternative strategy profiles $\widehat{m}$,*

$$S(m|h_t) + \epsilon \geq S(\widehat{m}|h_t).$$

We note that conditional payoffs $\gamma_i(m_i, m_{-i}|h_T)$ and conditional output $S(m|h_T)$ are scaled by $\frac{1}{N}$ even if $T$ is large. Hence deviation temptations at time $T$ close to $N$ may in fact be large if scaled by the remaining number of periods $N - T$, rather than $N$. Mailath et al. (2005) make this point and suggest an alternative solution concept, contemporaneous perfect $\epsilon$-equilibrium, which addresses this issue in infinite horizon discounted games. We extend our analysis to contemporaneous perfect $\epsilon$-equilibrium in Section 7.

For simplicity, we assume as in Section 4 that $y_t$ is observed by the principal. The adjustment of Section 5 continues to apply.

### 6.1 Target Properties

We now formulate target incentive properties that will ensure dynamically consistent implementation. In each period, the mechanism chooses a distribution $\mu_t \in \Delta(\mathcal{P}(I))$ defining a random decision group, and feasible transfers $\tau_{i,t}(a)$ given the chosen allocation $a$.

**Smoothed externalities.** In this private value setting, under efficient allocations, players have negative expected externalities on each other. We build a measure of realized externalities that ignores large deviations towards positive externalities while still correctly accounting for negative externalities.

Given $(\lambda_{i,t})_{t\geq 1}$, with $\lambda_{i,t} \in [0,1]$, we define our smoothed measure of externalities by

$$\Phi_{i,T} = \sum_{t=1}^{T} \lambda_{i,t}[Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t)].$$

We want a process $(\lambda_{i,t})_{i\in I, t\geq 1}$ such that for all $i \in I$,

$$\mathcal{R}_{1,i,T} \equiv \max_{T'\leq T} \left\{ -\sum_{t=T'}^{T} \lambda_{i,t}[Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t)] \right\} \leq o(T) \tag{14}$$

$$\mathcal{R}_{2,i,T} \equiv \max_{T'\leq T} \left\{ \sum_{t=T'}^{T} (1-\lambda_{i,t})[Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t)] \right\} \leq o(T). \tag{15}$$

Condition (14) ensures that measured externalities do not become large and positive. Condition (15) ensures that measured externalities correctly reflect negative externalities. The running maximum over start dates $T' \leq T$ ensure that these properties hold from the perspective of any history.

**Correct allocation and transfers.** For all agents $i \in I$, let

$$\mathcal{R}^{\tau}_{i,T} \equiv \Phi_{i,T} - \sum_{t=1}^{T} \tau_{i,t}(\mu_t) \quad \text{and} \quad \mathcal{R}^{I}_{T} \equiv \max_{T' \leq T} \sum_{t=T'}^{T} Y_t(I) - Y_t(\mu_t). \tag{16}$$

We want to pick $(\mu_t, \tau_{i,t})_{i\in I, t\geq 1}$ such that

$$|\mathcal{R}^{\tau}_{i,T}| = o(T) \tag{17}$$

$$\mathcal{R}^{I}_{T} \leq o(T). \tag{18}$$

These properties differ from original requirements (5) and (4) in that players need to pay their externality as measured by $(\Phi_{i,T})_{i\in I}$, and performance losses at time $T$ must be negligible compared to $T$ starting from any period $T' \leq T$. This ensures that the incentive properties of the pivot mechanism hold approximately from the perspective of any history.

## 6.2 An Explicit Mechanism

The standard approachability argument of Blackwell (1956) implies that conditions (14) and (15) can be satisfied by setting

$$\lambda_{i,T+1} = \frac{\mathcal{R}^{+}_{2,i,T}}{\mathcal{R}^{+}_{1,i,T} + \mathcal{R}^{+}_{2,i,T}}. \tag{19}$$

To see how (17) and (18) can be jointly satisfied, consider the vector of regrets $\mathcal{R}_T = (\mathcal{R}^{I}_{T}, \mathcal{R}^{\tau}_{i,T})_{i\in I}$. The approachability condition can be written as

$$\begin{aligned}
\left\langle \mathcal{R}^{+}_{T}, \mathcal{R}_{T+1} - \mathcal{R}_{T} \right\rangle = &\sum_{i\in I} \left[\mathcal{R}^{\tau}_{i,T}\right]^{+} \left[\lambda_{i,T+1}(Y_{-i,T+1}(I\backslash i) - Y_{-i,T+1}(\mu_{T+1})) - \tau_{i,T+1}(\mu_{T+1})\right] \\
&+ \left[\mathcal{R}^{I}_{T}\right]^{+} (Y_{T+1}(I) - Y_{T+1}(\mu_{T+1})) \\
= &\sum_{i\in I} \left[\lambda_{i,T+1}\mathcal{R}^{\tau}_{i,T}\right]^{+} [y_{i,T+1}(\mu_{T+1}) - y_{i,T+1}(I\backslash i)] - \sum_{i\in I} \left[\mathcal{R}^{\tau}_{i,T}\right]^{+} \tau_{i,T+1}(\mu_{T+1})
\end{aligned}$$

$$+\left[\mathcal{R}_T^I\right]^+ (Y_{T+1}(I) - Y_{T+1}(\mu_{T+1})) + \sum_{i\in I} \left[\lambda_{i,T+1}\mathcal{R}_{i,T}^\tau\right]^+ [Y_{T+1}(I\backslash i) - Y_{T+1}(\mu_{T+1})].$$

Note that $\lambda_{i,T+1}$ depends only on history up to period $T$. It follows that approachability condition

$$\left\langle \mathcal{R}_T^+, \mathcal{R}_{T+1} - \mathcal{R}_T \right\rangle \leq 0 \tag{20}$$

is satisfied by using the allocation rule $\mu_{T+1}$ such that for all $i \in I$,

$$\mu_{T+1}(I\backslash i) = \frac{[\lambda_{i,T+1}\mathcal{R}_{i,T}^\tau]^+}{[\mathcal{R}_T^I]^+ + \sum_{j\in I}[\lambda_{j,T+1}\mathcal{R}_{j,T}^\tau]^+} \tag{21}$$

and $\mu_{T+1}(I) \equiv 1 - \sum_{i\in I}\mu_{T+1}(I\setminus i)$, as well as setting transfers

$$\tau_{i,T+1}(a_{T+1}) = \begin{cases} y_{i,T+1}(a_{T+1}) & \text{if} \quad \mathcal{R}_{i,T}^\tau > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

**Proposition 4** *Consider the mechanism $(\mu_t, \tau_{i,t})_{i\in I, t\geq 1}$ defined by (21) and (22). For all $\epsilon > 0$, there exists $N_0$ such that for all $N > N_0$, truthful revelation is a perfect $\epsilon$-equilibrium and $\epsilon$-renegotiation proof.*

**Proof:** We begin by showing that processes $(\mu_t, \lambda_{i,t}, \tau_{i,t})_{i\in I, t\geq 1}$ ensure that target properties (14), (15), (17), (18) are satisfied.

Let us begin with (14) and (15). Pick any $i \in I$. Maximization over start dates $T'$ in the expression for $\mathcal{R}_{1,i,T}$ and $\mathcal{R}_{2,i,T}$, implies that

$$\mathcal{R}_{1,i,T+1} = \mathcal{R}_{1,i,T}^+ - \lambda_{i,T+1} \times (Y_{-i,T+1}(I\backslash i) - Y_{-i,T+1}(\mu_{T+1}))$$
$$\mathcal{R}_{2,i,T+1} = \mathcal{R}_{2,i,T}^+ + (1 - \lambda_{i,T+1}) \times (Y_{-i,T+1}(I\backslash i) - Y_{-i,T+1}(\mu_{T+1}))$$

In addition, observe that for all values $(R, \Delta) \in \mathbb{R}^2$, $R^+ \times (R^+ + \Delta - R) = R^+\Delta$. Altogether, this implies the following approachability condition:

$$\mathcal{R}^+_{1,i,T} \times (\mathcal{R}_{1,i,T+1} - \mathcal{R}_{1,i,T}) + \mathcal{R}^+_{2,i,T} \times (\mathcal{R}_{2,i,T+1} - \mathcal{R}_{2,i,T}) \leq \\ \left(-\mathcal{R}^+_{1,i,T}\lambda_{i,T+1} + (1-\lambda_{i,T+1})\mathcal{R}^+_{2,i,T}\right) \times [Y_{-i,T+1}(I \backslash i) - Y_{-i,T+1}(\mu_{T+1})] = 0.$$

Inequalities (14) and (15) follow directly from this approachability condition.

We now prove (17) and (18). The fact that $(\mu_t)_{t\geq 1}$ satisfies approachability condition (20) implies that $||\mathcal{R}^+_T|| = O(\sqrt{T})$. This implies (18), as well as $\mathcal{R}^\tau_{i,T} \leq O(\sqrt{T})$. To establish (17), we need to show that $\mathcal{R}^\tau_{i,T} \geq -O(\sqrt{T})$.

For any period $T$, define $\underline{T} = \max\{t < T | \tau_{i,t}(\mu_t) > 0\}$. By construction, this implies that $\Phi_{i,\underline{T}-1} - \sum_{t=1}^{\underline{T}-1} \tau_{i,t} \geq 0$. Since $\tau_{i,t} = 0$ for all $t > \underline{T}$, it follows that

$$\begin{aligned}\Phi_{i,T} - \sum_{t=1}^{T} \tau_{i,t} =& \Phi_{i,\underline{T}-1} - \sum_{t=1}^{\underline{T}-1} \tau_{i,t} - \tau_{i,\underline{T}} + \Phi_{i,T} - \Phi_{i,\underline{T}-1} \\ \geq & -O(\sqrt{T})\end{aligned}$$

where we used (14) to obtain a lower bound for $\Phi_{i,T} - \Phi_{i,\underline{T}-1}$. This establishes (17).

We now show that target properties (14), (15), (17), (18) imply that truthtelling is an $\epsilon$-Nash equilibrium from the perspective of every history $h_T$. The proof is very similar to that of Proposition 2. As in the proof of Proposition 2, target properties hold with $o(T)$ dependent only on $T$, $|I|$, and $y_{\max}$, and so they hold in expectation with the same $o(T)$ term for any messaging strategies.

Assume that $m_{-i} = v_{-i}$, i.e. that players other than $i$ are reporting truthfully. For any history $h_T$, we have that

$$\begin{aligned}\mathbb{E}_{m_i,v_{-i}}\left(\sum_{t=T}^{N} y_{i,t}(\mu_t) - \tau_{i,t}(\mu_t)\middle| h_T\right) &\leq \mathbb{E}_{m_i,v_{-i}}\left(\sum_{t=T}^{N} y_{i,t}(\mu_t) - [\Phi_{i,N} - \Phi_{i,T-1}]\middle| h_T\right) + o(N) \\ &\leq \mathbb{E}_{m_i,v_{-i}}\left(\sum_{t=T}^{N} y_{i,t}(\mu_t) - [Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t)]\middle| h_T\right) + o(N)\end{aligned}$$

$$\leq \mathbb{E}_{v_i,v_{-i}} \left( \sum_{t=T}^{N} Y_t(I) - Y_{-i,t}(I \backslash i) \middle| h_T \right) + o(N)$$

where the first and second steps respectively use (17) at $N$ and $T$ and (15). The $o(N)$ term is independent of $m_i$ and $h_T$.

In turn, truthtelling guarantees this upper-bound up to a negligible term:

$$\begin{aligned}
&\mathbb{E}_{v_i,v_{-i}} \left( \sum_{t=T}^{N} y_{i,t}(\mu_t) - \tau_{i,t}(\mu_t) \middle| h_T \right) \\
&\qquad \geq \mathbb{E}_{v_i,v_{-i}} \left( \sum_{t=T}^{N} y_{i,t}(\mu_t) + \sum_{t=T}^{N} \lambda_{i,t}[Y_{-i,t}(\mu_t) - Y_{-i,t}(I \backslash i)] \middle| h_T \right) - o(N) \\
&\qquad \geq \mathbb{E}_{v_i,v_{-i}} \left( \sum_{t=T}^{N} y_{i,t}(\mu_t) + \sum_{t=T}^{N} Y_{-i,t}(\mu_t) - Y_{-i,t}(I \backslash i) \middle| h_T \right) - o(N) \\
&\qquad \geq \mathbb{E}_{v_i,v_{-i}} \left( \sum_{t=T}^{N} Y_t(I) - Y_{-i,t}(I \backslash i) \middle| h_T \right) - o(N)
\end{aligned}$$

where we used (17) at $N$ and $T$, (18), and the fact that $\mathbb{E}[Y_{-i,t}(\mu_t) - Y_{-i,t}(I \backslash i)|h_t] \leq 0$. The $o(N)$ term is independent of $h_T$. The result follows for $N$ large enough that $o(N) \leq \epsilon N$.

Condition (18) implies $\epsilon$-renegotiation proofness. ■

The mechanism $(\mu_t, \tau_{i,t})_{i \in I, t \geq 1}$ described above exhibits both cautiousness and forgiveness. Consider the example of a player $i$ making erroneous reports of their values for $T_0$ periods and truthful reports from period $T_0$ on. In periods $T = 1, \ldots, T_0$, poor suggestions that are followed will cause regret $\mathcal{R}^{\tau}_{i,T}$ to grow. As a result player $i$ will be steadily excluded from the decision making process. Cautiousness limits the amount of debt player $i$ accrues for unpaid externalities. This implies that when the player starts making correct suggestions so that $\mathcal{R}^I_T = \max_{T' \leq T} \sum_{t=T'}^{T} Y_t(I) - Y_t(\mu_t)$ is large and positive, past debts are sufficiently small to ignore. Player $i$ is essentially forgiven, and reincluded in the decision-making process.[12] We emphasize the complementarity between cautiousness and forgiveness: if agents were not

[12] Hauser and Hopenhayn (2005) also finds a role for forgiveness in a model of trading favors: they help relax IC constraints after asymmetric histories of contributions.

cautiously excluded when they make erroneous reports, they could not be re-included swiftly once they start making correct reports.

# 7 Discounting

We now show how to adapt our approach to environments with an infinite horizon and discounting. Results for $T$ large are replaced by results for $\delta$ near 1.

Players now evaluate messaging strategies according to payoffs

$$\gamma_i(m_i, m_{-i}) = (1-\delta)\mathbb{E}\left(\sum_{t=1}^{\infty}\delta^{t-1}(y_{i,t}(a_t) - \tau_{i,t})\right)$$

The framework is otherwise the same as in Section 2. Given a mechanism, the game it induces is denoted by $\Gamma(\delta)$.

In this environment, we show that it is possible to implement truthtelling and approximately efficient allocations in contemporaneous perfect $\epsilon$-equilibrium (Mailath et al., 2005), rather than perfect $\epsilon$-equilibrium (Radner, 1980): this establishes a stronger form of time consistency by properly rescaling continuation payoffs so that current payoffs loom large at every history.

**Definition 2 (Mailath et al. (2005))** *A* contemporaneous perfect $\epsilon$-equilibrium *is a strategy profile $m = (m_i)_{i\in I}$ such that, after every history $h_T$ and for each player $i$,*

$$(1-\delta)\mathbb{E}_m\left(\sum_{t=T}^{\infty}\delta^{t-T}(y_{i,t}(\mu_t) - \tau_{i,t})\middle| h_T\right) \geq (1-\delta)\mathbb{E}_{(\tilde{m}_i, m_{-i})}\left(\sum_{t=T}^{\infty}\delta^{t-T}(y_{i,t}(\mu_t) - \tau_{i,t})\middle| h_T\right) - \epsilon$$

*for any alternative strategy $\tilde{m}_i$.*

*A strategy profile $m = (m_i)_{i\in I}$ is contemporaneous $\epsilon$-renegotiation proof if and only if for* *all histories $h_t$ and all alternative strategy profiles $\tilde{m}$,*

$$(1-\delta)\mathbb{E}_m\left(\sum_{t=T}^{\infty}\delta^{t-T}Y_t(a_t)\middle|h_T\right) \geq (1-\delta)\mathbb{E}_{\tilde{m}}\left(\sum_{t=T}^{\infty}\delta^{t-T}Y_t(a_t)\middle|h_T\right) - \epsilon$$

Continuation payoffs at $T$ are now evaluated from the perspective of time $T$, rather than time 1. Drawing on Schlag and Zapechelnyuk (2017), it turns out that a simple adjustment to the mechanism of Section 6 ensures that truthful reporting is both a contemporaneous perfect $\epsilon$-equilibrium, and contemporaneous $\epsilon$-renegotiation-proof.

Given arbitrary processes $(\lambda_{i,t}, \mu_t, \tau_{i,t})_{i\in I, t\in\mathbb{N}}$, we extend the definitions of regrets given by (14), (15) and (16) to include *backward discounting*:

$$\begin{aligned}
\mathcal{R}_{1,i,T} &\equiv \max_{T'\leq T}\left\{-\sum_{t=T'}^{T}\lambda_{i,t}\delta^{T-t}(Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t))\right\}\\
\mathcal{R}_{2,i,T} &\equiv \max_{T'\leq T}\left\{\sum_{t=T'}^{T}(1-\lambda_{i,t})\delta^{T-t}(Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t))\right\}\\
\mathcal{R}^{\tau}_{i,T} &\equiv \sum_{t=1}^{T}\delta^{T-t}(\lambda_{i,t}(Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t)) - \tau_{i,t}(\mu_t))\\
\mathcal{R}^{I}_{T} &\equiv \max_{T'\leq T}\left\{\sum_{t=T'}^{T}\delta^{T-t}(Y_t(I) - Y_t(\mu_t))\right\}.
\end{aligned}$$

Regrets defined in Section 6 are a special case of backward discounted regrets, with $\delta = 1$. Backward discounting ensures that early outcomes have a vanishing influence on regrets at later histories.

Processes $(\lambda_{i,t}, \mu_t, \tau_{i,t})_{i\in I, t\in\mathbb{N}}$ are defined according to (19), (21), and (22), using the backward discounted regrets above. This induces a dynamic mechanism for which the following holds.

**Proposition 5** *For all $\epsilon > 0$, there exists $\underline{\delta}$ such that for all $\delta \in (\underline{\delta}, 1)$, truthful revelation is a contemporaneous perfect $\epsilon$-equilibrium, and contemporaneous $\epsilon$-renegotiation proof.*

The proof is provided in Appendix B.2.

# 8 Discussion

**Summary.** This paper shows how to replicate the pivot mechanism in a prior-free way in dynamic environments with limited liability. The main insight is that by selectively excluding players unable to make their past externality payments, it is possible to ensure that in aggregate, players approximately pay for their externality and there is negligible allocation inefficiency. Under this mechanism, truthtelling is an $\epsilon$-Nash equilibrium for sufficiently long time horizons and the corresponding allocations are approximately efficient.

In addition the paper shows that by using a mix of cautiousness and forgiveness, it is possible to construct selective exclusion mechanisms that are dynamically consistent in the sense that truthful reporting is a perfect $\epsilon$-equilibrium and the corresponding allocation is approximately efficient from the perspective of any history.

We now discuss extensions tackled in Appendix A, as well as some of the many limits of our analysis.

**Limited observability.** A practical concern is that counterfactual returns need not be observable, even to the agents. We show in Appendix A that even in this environment, it is possible to enforce efficient allocation in $\epsilon$-Nash equilibrium.

**Hidden savings.** Our framework can potentially accommodate hidden savings, provided that the interest rate is zero. In that case, the marginal benefit of consuming income now, versus saving and consuming the resulting savings later is the same. Under the environment and the mechanism of Section 5, realized output is privately observed by each agent and reported as part of the mechanism. Under this mechanism, hidden consumption or hidden savings essentially correspond to misreporting output, and do not change the result that truthful reporting is an $\epsilon$-Nash equilibrium.

**Other VCG mechanisms.** In contexts without limited liability, there are many implementations of VCG that add a term depending only on the reports of players $j \neq i$ to the transfers of player $i$. However, the pivot mechanism is the only one that ensures individual rationality and non-negative transfers (Moulin, 1986). In our setting, the pivot mechanism is the only mechanism ensuring that: $\mathbb{E}[y_{i,t}(a_t) - \tau_{i,t}] \geq 0$, so that in expectation, players are able to make their required transfers using only their output; and $\tau_{i,t} \geq 0$, so that the principal does not need to subsidize agents. This implies that without relaxing constraints on feasible transfers we cannot use our approach to replicate VCG mechanisms other than the pivot.

**Common values.** Although our analysis makes use of the private value assumption, it is possible to accommodate mild forms of common values. One observation is that messages from agents' can no longer be interpreted as preferences, but instead must be interpreted as signals. Indeed, in common value environments, player $i$'s preferences depend on the signal submitted by other participants. The principal's policy then becomes a mapping from signals to allocations. The main difficulty lies in the fact that under common values, one player may have a positive externality on others, which means that the mechanism may involve negative transfers. Our approach can be extended under the limiting assumption that in aggregate, players have negative externalities on each other. This ensures that appropriate externality payments are feasible. More general extensions remain a challenge.

**Out-of-equilibrium payoff guarantees.** Another limitation of our analysis, in particular compared to Jackson and Sonnenschein (2007), is that we do not provide out-of-equilibrium payoff guarantees for players. Among other things, in Jackson and Sonnenschein (2007), out-of-equilibrium payoff guarantees imply that all equilibria sustain approximately efficient allocation.

In our setting, one difficulty is that even under a restriction to Markov strategies, the pivot mechanism may exhibit common value aspects out-of-equilibrium: one player may have a positive externality on others. Under strong assumptions ensuring negative externalities even off the equilibrium path, it is possible to establish meaningful payoff guarantees, however such strong assumptions are ultimately unsatisfactory. Providing satisfactory payoff-guarantees off of the equilibrium path remains an open challenge.

**Negative transfers.** With private values, transfers prescribed by the static VCG mechanism are positive. Because of this, our limited liability mechanism does not require negative transfers to approximate the static VCG mechanism. In a common value environment where agents can have a positive externality on each other, negative transfers can improve efficiency.

# Appendix

# A Unobserved Counterfactuals

We now assume that only outcomes for the decision $a_t$ that is actually taken in period $t$ are observed by the agents.

## A.1 Values Bounded Away From 0

In this section, we proceed under an additional assumption that values are bounded away from 0.

**Assumption A.1** *For all $i \in I, t \in \mathbb{N}, a \in A$,*

$$v_{i,t}(a) > \underline{v} > 0$$

To demonstrate the generality of the approach, we assume that the principal does not observe outputs, and so must rely on the agents' reports of both values and outputs. As in Section 5, a message $m_{i,t}$ consists of stated values $\overline{v}_{i,t}$ and stated outputs $\overline{y}_{i,t-1}$ for the action taken.

**Smoothed externalities and a mechanism.** As in Section 6, we will use smoothed externalities. Define $(a_t, \delta_{i,t}, \lambda_{i,t})_{i\in I, t\geq 1}$ recursively as follows. Given $(\lambda_{i,t})_{i\in I}$ and $G \subset I$, let

$$a_t^*(G) \in \arg\max_{a\in A} \sum_{i\in G} \lambda_{i,t}\overline{v}_{i,t}(a)$$

where $\overline{v}_{i,t}$ is agent $i$'s stated value to the mechanism. The allocation is set to

$$a_t = a_t^*(I). \tag{A.1}$$

Let

$$\overline{V}_t(a) = \sum_{i\in I} \lambda_{i,t}\overline{v}_{i,t}(a)$$
$$\forall i \in I, \quad \overline{V}_{-i,t}(a) = \sum_{j\neq i} \lambda_{j,t}\overline{v}_{j,t}(a)$$

Given $\delta_{i,t} \in \{0,1\}$, transfers are set to

$$\tau_{i,t} = \delta_{i,t}\overline{y}_{i,t}. \tag{A.2}$$

Regrets are,

$$\mathcal{R}^\tau_{i,T} = \sum_{t=1}^T \overline{V}_{-i,t}(a_t^*(I\backslash i)) - \overline{V}_{-i,t}(a_t^*(I)) - \tau_{i,t}$$
$$\mathcal{R}^v_{i,T} = \sum_{t=1}^T \overline{V}_{-i,t}(a_t^*(I\backslash i)) - \overline{V}_{-i,t}(a_t^*(I)) - \delta_{i,t}\lambda_{i,t}\overline{v}_{i,t}(a_t^*(I))$$
$$\mathcal{R}^A_{i,T} = \sum_{t=1}^T \left[\overline{y}_{i,t} - \lambda_{i,t}\overline{v}_{i,t}(a_t)\right] \mathbf{1}_{\mathcal{R}^v_{i,t-1}>0}$$
$$\mathcal{R}^B_{i,T} = \sum_{t=1}^T \left[\overline{y}_{i,t} - \lambda_{i,t}\overline{v}_{i,t}(a_t)\right] \mathbf{1}_{\mathcal{R}^v_{i,t-1}\leq 0}$$

Finally, set

$$\delta_{i,T+1} = \begin{cases} 1 \text{ if } \mathcal{R}^v_{i,T} > 0 \\ 0 \text{ if } \mathcal{R}^v_{i,T} \leq 0 \end{cases}$$

$$\lambda_{i,T+1} = \begin{cases} 1 \text{ if } \mathcal{R}^v_{i,T} > 0 \text{ and } \mathcal{R}^A_{i,T} \geq 0 \\ 0 \text{ if } \mathcal{R}^v_{i,T} > 0 \text{ and } \mathcal{R}^A_{i,T} < 0 \\ 1 \text{ if } \mathcal{R}^v_{i,T} \leq 0 \text{ and } \mathcal{R}^B_{i,T} \geq 0 \\ 0 \text{ if } \mathcal{R}^v_{i,T} \leq 0 \text{ and } \mathcal{R}^B_{i,T} < 0 \end{cases}$$

**Lemma A.1** *Consider the mechanism $(a_t, (\tau_{i,t})_{i\in I})_{t\geq 1}$ defined by (A.1) and (A.2). For all strategies and $i \in I$*

$$\mathcal{R}^A_{i,T} \geq o(T) \tag{A.3}$$
$$\mathcal{R}^B_{i,T} \geq o(T) \tag{A.4}$$
$$\mathcal{R}^v_{i,T} = o(T) \tag{A.5}$$
$$\mathcal{R}^\tau_{i,T} \leq o(T) \tag{A.6}$$

*where $o(T)$ depends only on $T$, $|I|$, and $y_{\max}$. For any strategy profile $m$ such that $i$ reports truthfully,*

$$\mathbb{E}_m\left(\mathcal{R}^A_{i,T}\right) \leq o(T) \tag{A.7}$$
$$\mathbb{E}_m\left(\mathcal{R}^B_{i,T}\right) \leq o(T) \tag{A.8}$$
$$\mathbb{E}_m\left(\sum_{t=1}^{T}(1-\lambda_{i,t})\right) \leq o(T) \tag{A.9}$$
$$\mathbb{E}_m\left(\mathcal{R}^\tau_{i,T}\right) \geq o(T) \tag{A.10}$$

*where $o(T)$ depends only on $T$, $|I|$, and $y_{\max}$.*

**Proof:** We begin by proving (A.3). Denote $\Delta\mathcal{R}^A_{i,T+1} = [\overline{y}_{i,T+1} - \lambda_{i,T+1}\overline{v}_{i,T+1}(a_{T+1})]\mathbf{1}_{\mathcal{R}^v_{i,T}>0}$. Assume that $\mathcal{R}^A_{i,T} < 0$. Then, $\Delta\mathcal{R}^A_{i,T+1} = \mathbf{1}_{\mathcal{R}^v_{i,T}>0}\overline{y}_{i,T+1} \geq 0$ and (A.3) follows. An analogous proof leads to (A.4).

Next we prove (A.5). Let

$$\Delta\mathcal{R}^v_{i,T+1} = \overbrace{\overline{V}_{-i,T+1}(a^*_{T+1}(I\backslash i)) - \overline{V}_{-i,T+1}(a^*_{T+1}(I))}^{\geq 0} - \delta_{i,T+1}\lambda_{i,T+1}\overline{v}_{i,T+1}(a^*_{T+1}(I))$$

If $\mathcal{R}^v_{i,T} \leq 0$, then $\delta_{i,T+1} = 0$, which implies that $\Delta\mathcal{R}^v_{i,T+1} \geq 0$. This implies that $\mathcal{R}^v_{i,T} \geq o(T)$.

If instead $\mathcal{R}^v_{i,T} > 0$, then $\delta_{i,T+1} = 1$. Observe then

$$\begin{aligned}\Delta\mathcal{R}^v_{i,T+1} = \; & \overline{V}_{-i,T+1}(a^*_{T+1}(I\backslash i)) - \overline{V}_{-i,T+1}(a^*_{T+1}(I)) - \lambda_{i,T+1}\overline{v}_{i,T+1}(a^*_{T+1}(I)) \\ = \; & \overline{V}_{T+1}(a^*_{T+1}(I\backslash i)) - \overline{V}_{T+1}(a^*_{T+1}(I)) \\ & + \lambda_{i,T+1}\left(\overline{v}_{i,T+1}(a^*_{T+1}(I)) - \overline{v}_{i,T+1}(a^*_{T+1}(I\backslash i))\right) - \lambda_{i,T+1}\overline{v}_{i,T+1}(a^*_{T+1}(I)) \\ = \; & \overbrace{\overline{V}_{T+1}(a^*_{T+1}(I\backslash i)) - \overline{V}_{T+1}(a^*_{T+1}(I))}^{\leq 0} - \lambda_{i,T+1}\overline{v}_{i,T+1}(a^*_{T+1}(I\backslash i)).\end{aligned}$$

In this case, $\Delta\mathcal{R}^v_{i,T+1} \leq 0$, and so $\mathcal{R}^v_{i,T} \leq o(T)$. Combining with our finding that $\mathcal{R}^v_{i,T} \geq o(T)$, we have

$$\mathcal{R}^v_{i,T} = o(T).$$

Next we prove (A.6). Observe,

$$\begin{aligned}\mathcal{R}^\tau_{i,T} &= \mathcal{R}^v_{i,T} - \sum_{t=1}^{T}\left[\overline{y}_{i,t} - \lambda_{i,t}\overline{v}_{i,t}(a_t)\right]\mathbf{1}_{\mathcal{R}^v_{i,t-1}>0} \\ &= \mathcal{R}^v_{i,T} - \mathcal{R}^A_{i,T} \\ &\leq o(T)\end{aligned}$$

where the third line follows by (A.3) and (A.5).

For the purpose of proving the remaining results, denote by $m^*_i$ truthful revelation for player $i$ and $m_{-i}$ an arbitrary strategy profile for $-i$.

Next we prove (A.7). Define the positive part of regret, $\mathcal{R}^{A,+}_{i,T} \equiv \max\{0, \mathcal{R}^A_{i,T}\}$. If $\mathcal{R}^A_{i,T} \geq 0$ and $\mathcal{R}^v_{i,T} > 0$, then $\Delta\mathcal{R}^A_{i,T+1} = \overline{y}_{i,T+1} - \overline{v}_{i,T+1}(a_{T+1})$. If $\mathcal{R}^A_{i,T} \geq 0$ but $\mathcal{R}^v_{i,T} \leq 0$, then $\Delta\mathcal{R}^A_{i,T+1} = 0$. As a result,

$$\mathbb{E}_{m^*_i,m_{-i}}\left(\mathcal{R}^{A,+}_{i,T} \times \Delta\mathcal{R}^A_{i,T+1}\right) = 0.$$

Using this fact, it follows that

$$\begin{aligned}\mathbb{E}_{m^*_i,m_{-i}}\left(\left[\mathcal{R}^{A,+}_{i,T+1}\right]^2\right) &\leq \mathbb{E}_{m^*_i,m_{-i}}\left(\left[\mathcal{R}^{A,+}_{i,T}\right]^2 + 2\times\mathcal{R}^{A,+}_{i,T}\times\Delta\mathcal{R}^A_{i,T+1} + \left[\Delta\mathcal{R}^A_{i,T+1}\right]^2\right) \\ &= \mathbb{E}_{m^*_i,m_{-i}}\left(\left[\mathcal{R}^{A,+}_{i,T}\right]^2\right) + \mathbb{E}_{m^*_i,m_{-i}}\left(\left[\Delta\mathcal{R}^A_{i,T+1}\right]^2\right) \\ &\leq \mathbb{E}_{m^*_i,m_{-i}}\left(\left[\mathcal{R}^{A,+}_{i,T}\right]^2\right) + y^2_{\max}\end{aligned}$$

$$\leq (T+1)y_{\max}^2.$$

Jensen's inequality implies $\mathbb{E}_{m_i^*,m_{-i}}\left(\mathcal{R}_{i,T+1}^{A,+}\right) \leq \sqrt{\mathbb{E}_{m_i^*,m_{-i}}\left(\left[\mathcal{R}_{i,T+1}^{A,+}\right]^2\right)} \leq y_{\max}\sqrt{T+1}$, which yields (A.7). Similar steps lead to (A.8).

Next we prove (A.9). By summing (A.7) and (A.8), we have

$$\begin{aligned}
&\mathbb{E}_{m_i^*,m_{-i}}\left(\sum_{t=1}^{T}\overline{y}_{i,t} - \lambda_{i,t}\overline{v}_{i,t}(a_t)\right) \leq o(T) \\
\Rightarrow\quad &\mathbb{E}_{m_i^*,m_{-i}}\left(\sum_{t=1}^{T}(1-\lambda_{i,t})v_{i,t}(a_t)\right) \leq o(T) \\
\Rightarrow\quad &\mathbb{E}_{m_i^*,m_{-i}}\left(\underline{v}\sum_{t=1}^{T}(1-\lambda_{i,t})\right) \leq o(T) \\
\Rightarrow\quad &\mathbb{E}_{m_i^*,m_{-i}}\left(\sum_{t=1}^{T}(1-\lambda_{i,t})\right) \leq o(T)
\end{aligned}$$

where the last line follows by Assumption A.1.

Finally, we prove (A.10). As in the proof of (A.6), observe that

$$\begin{aligned}
\mathcal{R}_{i,T}^{\tau} &= \mathcal{R}_{i,T}^{v} - \mathcal{R}_{i,T}^{A} \\
\Rightarrow \mathbb{E}_{m_i^*,m_{-i}}\left(\mathcal{R}_{i,T}^{\tau}\right) &= \mathbb{E}_{m_i^*,m_{-i}}\left(\mathcal{R}_{i,T}^{v}\right) - \mathbb{E}_{m_i^*,m_{-i}}\left(\mathcal{R}_{i,T}^{A}\right) \\
&\geq o(T)
\end{aligned}$$

where the last line follows by (A.5) and (A.7). ■

Define

$$\overline{a}_t^*(G) \in \arg\max_{a\in A}\sum_{i\in G}\overline{v}_{i,t}(a),$$

i.e. an action that maximizes the unweighted sum of stated values.

**Lemma A.2** *Consider the mechanism $(a_t,(\tau_{i,t})_{i\in I})_{t\geq 1}$ defined by (A.1) and (A.2). For any strategy profile $m$ such that agents $i \in G$ report truthfully,*

$$\sum_{t=1}^{T}\mathbb{E}_m\left(\sum_{i\in G}v_{i,t}(\overline{a}_t^*(G)) - v_{i,t}(a_t^*(G))\right) = o(T) \tag{A.11}$$

*where $o(T)$ depends only on $T$, $|I|$ and $y_{\max}$.*

**Proof:** Let $m_G^*$ denote truthful revelation by agents $i \in G$ and $m_{-G}$ an arbitrary strategy profile for agents $i \in I \backslash G$.

Let $E_t$ be the event $\left\{ \prod_{i \in G} \lambda_{i,t} = 0 \right\}$. Then,

$$\sum_{t=1}^{T} \mathbb{E}_{m_G^*, m_{-G}} \left( \mathbf{1}_{E_t} \right) \leq \sum_{t=1}^{T} \mathbb{E}_{m_G^*, m_{-G}} \left( \sum_{i \in G} (1 - \lambda_{i,t}) \right) \leq o(T) \tag{A.12}$$

where $\mathbf{1}_{E_t}$ is an indicator for event $E_t$ and the second inequality follows by (A.9) for $i \in G$.

Let $E_t'$ be the event $\left\{ \sum_{i \in G} \overline{v}_{i,t}(\overline{a}_t^*(G)) \neq \sum_{i \in G} \overline{v}_{i,t}(a_t^*(G)) \right\}$. Observe that $E_t' \subseteq E_t$, so

$$\sum_{t=1}^{T} \mathbb{E}_{m_G^*, m_{-G}} \left( \mathbf{1}_{E_t'} \right) \leq \sum_{t=1}^{T} \mathbb{E}_{m_G^*, m_{-G}} \left( \mathbf{1}_{E_t} \right) \leq o(T)$$

where the second inequality follows by (A.12). Then,

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{E}_{m_G^*, m_{-G}} \left( \sum_{i \in G} v_{i,t}(\overline{a}_t^*(G)) - v_{i,t}(a_t^*(G)) \right) &= \sum_{t=1}^{T} \mathbb{E}_{m_G^*, m_{-G}} \left( \sum_{i \in G} \overline{v}_{i,t}(\overline{a}_t^*(G)) - \overline{v}_{i,t}(a_t^*(G)) \right) \\
&\leq \sum_{t=1}^{T} \mathbb{E}_{m_G^*, m_{-G}} \left( \mathbf{1}_{E_t'} \right) |I| y_{\max} \\
&\leq o(T).
\end{aligned}$$

Finally,

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{E}_{m_G^*, m_{-G}} \left( \sum_{i \in G} v_{i,t}(\overline{a}_t^*(G)) - v_{i,t}(a_t^*(G)) \right) &= \sum_{t=1}^{T} \mathbb{E}_{m_G^*, m_{-G}} \left( \sum_{i \in G} \overline{v}_{i,t}(\overline{a}_t^*(G)) - \overline{v}_{i,t}(a_t^*(G)) \right) \\
&\geq 0.
\end{aligned}$$

where the second line follows from the definition of $\overline{a}_t^*(G)$ as maximizing $\sum_{i \in G} \overline{v}_{i,t}(a)$. ■

A corollary of these lemmas is an analogue to Proposition 2.

**Proposition A.1** *For all $\epsilon > 0$, there exists $N_0$ such that for all $N > N_0$, truthful revelation is an $\epsilon$-Nash equilibrium of the game induced by the mechanism $(a_t, (\tau_{i,t})_{i \in I})_{t \geq 1}$ defined by*

(A.1) *and* (A.2). *Under truthful reporting, the allocation approaches efficiency as the horizon* $N$ *gets large.*

**Proof:** For any reporting strategy of $i$, $m_i$, and under truthful reporting of $-i$, $m^*_{-i}$,

$$\begin{aligned}
\mathbb{E}_{m_i,m^*_{-i}}\left(\sum_{t=1}^{N} y_{i,t} - \tau_{i,t}\right) &\leq \mathbb{E}_{m_i,m^*_{-i}}\left(\sum_{t=1}^{N} y_{i,t} + \sum_{j\neq i} \lambda_{j,t} v_{j,t}(a^*_t(I)) - \sum_{j\neq i} \lambda_{j,t} v_{j,t}(a^*_t(I\backslash i))\right) + o(N) \\
&\leq \mathbb{E}_{m_i,m^*_{-i}}\left(\sum_{t=1}^{N} y_{i,t} + \sum_{j\neq i} v_{j,t}(a^*_t(I)) - \sum_{j\neq i} v_{j,t}(a^*_t(I\backslash i))\right) + o(N) \\
&\leq \mathbb{E}_{m^*}\left(\sum_{t=1}^{N} \sum_{j\in I} v_{j,t}(\overline{a}^*_t(I)) - \sum_{j\neq i} v_{j,t}(\overline{a}^*_t(I\backslash i))\right) + o(N)
\end{aligned}$$

where the first line follows by (A.6), the second follows by (A.9) for all $j \neq i$, and $o(N)$ does not depend on $m_i$. The third line follows by applying (A.11) to $G = I, I\backslash i$.

In turn,

$$\begin{aligned}
\mathbb{E}_{m^*}\left(\sum_{t=1}^{N} y_{i,t} - \tau_{i,t}\right) &\geq \mathbb{E}_{m^*}\left(\sum_{t=1}^{N} y_{i,t} + \sum_{j\neq i} \lambda_{j,t} v_{j,t}(a^*_t(I)) - \sum_{j\neq i} \lambda_{j,t} v_{j,t}(a^*_t(I\backslash i))\right) - o(N) \\
&\geq \mathbb{E}_{m^*}\left(\sum_{t=1}^{N} y_{i,t} + \sum_{j\neq i} v_{j,t}(a^*_t(I)) - \sum_{j\neq i} v_{j,t}(a^*_t(I\backslash i))\right) - o(N) \\
&\geq \mathbb{E}_{m^*}\left(\sum_{t=1}^{N} \sum_{j\in I} v_{j,t}(\overline{a}^*_t(I)) - \sum_{j\neq i} v_{j,t}(\overline{a}^*_t(I\backslash i))\right) - o(N)
\end{aligned}$$

where the first line follows by (A.10), the second follows by (A.9) for all $j \neq i$, and the third line follows by applying (A.11) to $G = I, I\backslash i$.

The result that truth-telling is an $\epsilon$-Nash equilibrium follows by choosing $N$ sufficiently large so that $o(N) < \epsilon N$. Under truthful revelation, (A.11) implies that the allocation approaches efficiency as $N$ grows large. ■

In the next section, we establish slightly worse results in environments that violate Assumption A.1, by using random experimentation.

## A.2 General Values and Experimentation

We maintain the assumption that only outcomes for the decision $a_t$ that is actually taken in period $t$ are observed by the agents, but relax Assumption A.1. In particular, we revert to the assumption in the main text that $v_{i,t} \in [0, y_{\max}]$. For simplicity, we assume that the principal also observes the outcomes for the decision $a_t$ that is actually taken in period $t$.

The literature on online bandits (see Cesa-Bianchi and Lugosi, 2006, for an overview) shows that it is possible to minimize regrets with large probability by estimating regrets through experimentation. We outline an extension of our approach using such a strategy.[13]

**Estimated outputs.** Since actual outputs are not observed, we use instead estimated outputs. Given a full support distribution $\mu_t \in \Delta(\{I, I\backslash i | i \in I\})$ for random decision group $I_t$, estimated outputs $\widehat{y}_t$ are defined by

$$\forall G \in \mathsf{supp}\mu_t, \quad \widehat{y}_{i,t}(G) = \frac{1}{\mu_t(G)} \mathbf{1}_{I_t = G} \times y_{i,t}(I_t).$$

Note that $\widehat{y}_{i,t}(G)$ is an unbiased estimator of $y_{i,t}(G)$ such that

$$|\widehat{y}_{i,t}(G) - y_{i,t}(I\backslash i)| \leq \frac{y_{\max}}{\mu_t(G)}.$$

We denote by $\widehat{Y}_t$ and $\widehat{Y}_{-i,t}$ estimated counterparts of $Y_t$ and $Y_{-i,t}$.

**Estimated regrets.** Estimated outputs allow us to define estimated incentive and efficiency regrets as follows:

$$\text{(incentive alignment)} \qquad \forall i, T, \quad \widehat{\mathcal{R}}^{\tau}_{i,T} = \sum_{t=1}^{T} \widehat{Y}_{-i,t}(I\backslash i) - \widehat{Y}_{-i,t}(\mu_t) - \widehat{\tau}_{i,t} \leq o(T)$$

$$\text{(efficient allocation)} \qquad \forall T, \quad \widehat{\mathcal{R}}^{I}_{T} = \sum_{t=1}^{T} \widehat{Y}_t(I) - \widehat{Y}_t(\mu_t) \leq o(T)$$

The corresponding true regrets are $\mathcal{R}^{\tau}_{i,T} = \sum_{t=1}^{T} Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t) - \tau_{i,t}$ and $\mathcal{R}^{I}_{T} = \sum_{t=1}^{T} Y_t(I) - Y_t(\mu_t)$, as in Section 4.

**A mechanism.** The main idea is to ensure that allocation rule $(\mu_t)_{t\geq 1}$ explores at a rate that allows to identify counterfactual regrets without letting regrets grow large.

For all $T > 0$ let the distribution $\mu_{T+1}$ over decision groups in $\{I, I \setminus i \,|\, i \in I\}$ take the form

$$\mu_{T+1} = (1 - \nu_{T+1})\mu^*_{T+1} + \nu_{T+1}\mu_0$$

[13] Note that we do not optimize the exponents governing the speed at which regrets vanish.

where $\nu_t = t^{-a}$, with $a \in \left(0, \frac{1}{2}\right)$, $\mu_0$ is the uniform distribution over $\{I, I \setminus i \,|\, i \in I\}$, and

$$\forall i \in I, \quad \mu^*_{T+1}(I \backslash i) \equiv \frac{\left[\widehat{R}^{\tau}_{i,T}\right]^+}{\left[\widehat{R}^{I}_{T}\right]^+ + \sum_{j \in I}\left[\widehat{R}^{\tau}_{j,T}\right]^+}.$$

Transfers are defined as follows:

$$\widehat{\tau}_{i,T+1}(a_{T+1}) = \begin{cases} y_{i,T+1}(a_{T+1}) & \text{if} \quad \widehat{\mathcal{R}}^{\tau}_{i,T} > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The main technical step of the analysis, which is standard in the literature on online bandits shows that controlling estimated regrets also allows us to control expected true regrets.

**Lemma A.3** *For all $\eta > 0$, there exists $N_0$ such that for all $T > N_0$ and for all messaging strategies $m = (m_{i,t})_{i \in I, t \geq 1}$*

$$\mathbb{E}\mathcal{R}^{\tau}_{i,T} \leq \eta T \tag{A.13}$$

$$\mathbb{E}\mathcal{R}^{I}_{T} \leq \eta T. \tag{A.14}$$

**Proof:** Consider the vector of true regrets $\mathcal{R}_T = (\mathcal{R}^{\tau}_{i,T}, \mathcal{R}^{I}_{T})_{i \in I}$, and the vector of estimated regrets $\widehat{\mathcal{R}}_T = (\widehat{\mathcal{R}}^{\tau}_{i,T}, \widehat{\mathcal{R}}^{I}_{T})_{i \in I}$. We have that

$$\mathbb{E}||\mathcal{R}^+_{T+1}||^2 \leq \mathbb{E}||\mathcal{R}^+_{T}||^2 + 2\mathbb{E}\left\langle \mathcal{R}^+_T, \mathcal{R}_{T+1} - \mathcal{R}_T \right\rangle + \mathbb{E}||\mathcal{R}_{T+1} - \mathcal{R}_T||^2. \tag{A.15}$$

The last term is bounded by a constant independent of $T$: $\mathbb{E}||\mathcal{R}_{T+1} - \mathcal{R}_T||^2 \leq (|I|+1)^3 y^2_{\max}$. We also have that

$$\begin{aligned}
\mathbb{E}_{\mu_{T+1}}\left\langle \mathcal{R}^+_T, \mathcal{R}_{T+1} - \mathcal{R}_T \right\rangle =& \mathbb{E}\left\langle \mathcal{R}^+_T, \widehat{\mathcal{R}}_{T+1} - \widehat{\mathcal{R}}_T \right\rangle \\
\leq& \nu_{T+1}\mathbb{E}_{\mu_0}\left\langle \mathcal{R}^+_T, \widehat{\mathcal{R}}_{T+1} - \widehat{\mathcal{R}}_T \right\rangle + (1-\nu_{T+1})\mathbb{E}_{\mu^*_{T+1}}\left\langle \mathcal{R}^+_T, \widehat{\mathcal{R}}_{T+1} - \widehat{\mathcal{R}}_T \right\rangle \\
\leq& \nu_{T+1} T (|I|+1)^3 y^2_{\max} + (1-\nu_{T+1})\mathbb{E}_{\mu^*_{T+1}}\left\langle \mathcal{R}^+_T, \widehat{\mathcal{R}}_{T+1} - \widehat{\mathcal{R}}_T \right\rangle \\
\leq& \nu_{T+1} T (|I|+1)^3 y^2_{\max} + (1-\nu_{T+1})\mathbb{E}_{\mu^*_{T+1}}\left\langle \widehat{\mathcal{R}}^+_T, \widehat{\mathcal{R}}_{T+1} - \widehat{\mathcal{R}}_T \right\rangle \\
&\quad + (1-\nu_{T+1}) y_{\max}\mathbb{E}_{\mu^*_{T+1}}||\mathcal{R}^+_T - \widehat{\mathcal{R}}^+_T||_1 \\
\leq& \nu_{T+1} T (|I|+1)^3 y^2_{\max} + y_{\max}\mathbb{E}_{\mu^*_{T+1}}||\mathcal{R}^+_T - \widehat{\mathcal{R}}^+_T||_1
\end{aligned}$$

where we used the fact that $\mathbb{E}_{\mu^*_{T+1}} \left\langle \widehat{\mathcal{R}}^+_T, \widehat{\mathcal{R}}_{T+1} - \widehat{\mathcal{R}}_T \right\rangle = 0.$[14] In turn we have that

$$||\mathcal{R}^+_T - \widehat{\mathcal{R}}^+_T||_1 \leq ||\mathcal{R}_T - \widehat{\mathcal{R}}_T||_1 = \left| \sum_{t=1}^T Y_t - \widehat{Y}_t \right| + \sum_{i \in I} \left| \sum_{t=1}^T Y_{-i,t} - \widehat{Y}_{-i,t} \right|.$$

Terms $Y_t - \widehat{Y}_t$ and $Y_{-i,t} - \widehat{Y}_{-i,t}$ are both martingale increments $\Delta_t$ such that $|\Delta_t| \leq y_{\max}(|I|+1)/\nu_t$. It follows from the Azuma-Hoeffding theorem (see for instance Cesa-Bianchi and Lugosi, 2006) that for any $s > 0$, such sums of martingale increments satisfy

$$\textsf{prob}\left( \left| \sum_{t=1}^T \Delta_t \right| \geq s \right) \leq 2 \exp\left( \frac{-2s^2}{y^2_{\max}(|I|+1)^2 \sum_{t=1}^T \nu_t^{-2}} \right).$$

This implies that

$$\begin{aligned}
\mathbb{E} \left| \sum_{t=1}^T \Delta_t \right| \leq & \int_{s>0} \textsf{prob}\left( \left| \sum_{t=1}^T \Delta_t \right| > s \right) \mathrm{d}s \\
\leq & 2 \int_{s>0} \exp\left( \frac{-2s^2}{y^2_{\max}(|I|+1)^2 \sum_{t=1}^T \nu_t^{-2}} \right) \mathrm{d}s \\
\leq & 2\sqrt{\pi}(|I|+1) y_{\max} \sqrt{\sum_{t=1}^T \nu_t^{-2}}.
\end{aligned}$$

Using the fact that

$$\sum_{t=1}^T \nu_t^{-2} \leq \int_0^{T+1} s^{2a} \, \mathrm{d}s \leq \frac{1}{1+2a}(T+1)^{1+2a},$$

we obtain that there exists a constant $K_0$ depending only on $|I|$ such that

$$\mathbb{E}||\mathcal{R}^+_T - \widehat{\mathcal{R}}^+_T||_1 \leq K_0 (T+1)^{\frac{1}{2}+a}.$$

Replacing iteratively in (A.15) yields that there exists a constant $K_1$ such that

$$\begin{aligned}
\mathbb{E}||\mathcal{R}^+_{T+1}||^2 \leq & K_1 \left[ \sum_{t=1}^{T+1} t^{1-a} + (t+1)^{\frac{1}{2}+a} \right] \\
\leq & K_1 \left[ (T+2)^{2-a} + (T+2)^{\frac{3}{2}+a} \right],
\end{aligned}$$

[14] For all $x \in \mathbb{R}^n$, $||x||_1$ denotes the $L_1$ norm $||x||_1 \equiv \sum_{k=1}^n |x_k|$.

which implies that $\mathbb{E}||\mathcal{R}_T^+|| \leq K(T+2)^{\max\{1-\frac{a}{2},\frac{3}{4}+\frac{a}{2}\}}$. Since $a \in \left(0, \frac{1}{2}\right)$ this implies that for $T$ sufficiently large $\mathbb{E}||\mathcal{R}_T^+|| \leq \eta T$, which concludes the proof. ■

A corollary of this Lemma is an analogue to Proposition 2.

**Proposition A.2** *For all $\epsilon > 0$, there exists $N_0$ such that for all $N > N_0$, truthful revelation $(m_{i,t})_{i\in I,t\geq 1} = (v_{i,t})_{i\in I,t\geq 1}$ is an $\epsilon$-Nash equilibrium of $\Gamma$.*

The proof is identical to that of Proposition 2.

# B Other Proofs

## B.1 Proof of Proposition 3

Let $m_i^*$ denote the truthtelling strategy. Assume that $m_{-i} = m_{-i}^*$.

As in the proof of Proposition 2, the target properties (12) and (13) hold with $o(T)$ dependent only on $T$, $|I|$ and $y_{\max}$, which implies that they hold in expectation with the same $o(T)$ term for any messaging strategies.

By (12), we have that for any messaging strategy $m_i$,

$$\begin{aligned}\mathbb{E}_{m_i,m_{-i}^*}\left[\sum_{t=1}^{N} y_{i,t}(\mu_t) - \tau_{i,t}(\mu_t)\right] &\leq \mathbb{E}_{m_i,m_{-i}^*}\left[\sum_{t=1}^{N} Y_t(\mu_t) - Y_{-i,t}(I\backslash i)\right] + o(N)\\ &\leq \mathbb{E}_{m_i^*,m_{-i}^*}\left[\sum_{t=1}^{N} Y_t(I) - Y_{-i,t}(I\backslash i)\right] + o(N).\end{aligned}$$

where $o(N)$ is independent of $m_i$.

In turn, as in the proof of Proposition 2, the fact that $\mathbb{E}[Y_{-i,t}(\mu_t) - Y_{-i,t}(I \setminus i)] \leq 0$ and condition (13) imply that under truthtelling player $i$ can achieve a payoff

$$\begin{aligned}\mathbb{E}_{m_i^*,m_{-i}^*}\left[\sum_{t=1}^{N} y_{i,t}(\mu_t) - \tau_{i,t}(\mu_t)\right] &\geq \mathbb{E}_{m_i^*,m_{-i}^*}\left[\sum_{t=1}^{N} Y_t(\mu_t) - Y_{-i,t}(I\backslash i)\right] - o(N)\\ &\geq \mathbb{E}_{m_i^*,m_{-i}^*}\left[\sum_{t=1}^{N} Y_t(I) - Y_{-i,t}(I\backslash i)\right] - o(N).\end{aligned}$$

This implies that truthtelling is an $\epsilon$-best response for $N$ large enough.

## B.2 Proof of Proposition 5

**Target properties.** Take as given processes $(\lambda_{i,t}, \mu_t, \tau_{i,t})_{i \in I, t \in \mathbb{N}}$ defined in Section 7. For all $i \in I$, $T \in \mathbb{N}$, we define the following present values of future marginal regrets:

$$\mathcal{P}_{1,i,T} \equiv -\sum_{t=T}^{\infty} \delta^{t-T} \lambda_{i,t} \left[ Y_{-i,t}(I \setminus i) - Y_{-i,t}(\mu_t) \right]$$

$$\mathcal{P}_{2,i,T} \equiv \sum_{t=T}^{\infty} \delta^{t-T} (1 - \lambda_{i,t}) \left[ Y_{-i,t}(I \setminus i) - Y_{-i,t}(\mu_t) \right]$$

$$\mathcal{P}_T^I \equiv \sum_{t=T}^{\infty} \delta^{t-T} (Y_t(I) - Y_t(\mu_t))$$

$$\mathcal{P}_{i,T}^{\tau} \equiv \sum_{t=T}^{\infty} \delta^{t-T} (\lambda_{i,t}(Y_{-i,t}(I \setminus i) - Y_{-i,t}(\mu_t)) - \tau_{i,t}(\mu_t))$$

Our goal is to show that for all $i \in I$

$$\sup_{T \in \mathbb{N}} |\mathcal{P}_{i,T}^{\tau}| \leq o\left(\frac{1}{1-\delta}\right) \tag{B.16}$$

$$\sup_{T \in \mathbb{N}} \left[\mathcal{P}_T^I\right]^+ \leq o\left(\frac{1}{1-\delta}\right) \tag{B.17}$$

$$\sup_{T \in \mathbb{N}} \left[\mathcal{P}_{1,i,T}\right]^+ \leq o\left(\frac{1}{1-\delta}\right) \tag{B.18}$$

$$\sup_{T \in \mathbb{N}} \left[\mathcal{P}_{2,i,T}\right]^+ \leq o\left(\frac{1}{1-\delta}\right) \tag{B.19}$$

The difficulty is that these present values are forward looking, and involve marginal regrets that have not been observed at the time of decision making.

**Preliminary results.** An argument by Schlag and Zapechelnyuk (2017) allows us to exhibit strategies keeping forward-looking regrets small. It allows us to link backward-discounting regrets to present values. Let $(\Delta_t)_{t \in \mathbb{N}}$ denote an arbitrary sequence of marginal regrets.

**Lemma B.4** *Fix any $\Delta \in \mathbb{R}_+$ and suppose that for all $(\Delta_t)_{t \in \mathbb{N}}$ with $\sup_{t \in \mathbb{N}} |\Delta_t| \leq \Delta$,*

$$\sup_{T \in \mathbb{N}} \left| \sum_{t=1}^{T} \delta^{T-t} \Delta_t \right| \leq K\left(\frac{1}{\sqrt{1-\delta}}\right) \tag{B.20}$$

*for all* $\delta \in (0,1)$ *with constant* $K$ *independent of* $\delta$ *and* $(\Delta_t)_{t\in\mathbb{N}}$. *Then for all* $(\Delta_t)_{t\in\mathbb{N}}$ *such that* $\sup_{t\in\mathbb{N}} |\Delta_t| \le \Delta$,

$$\sup_{T\in\mathbb{N}} \left| \sum_{t=T}^{\infty} \delta^{t-T}\Delta_t \right| \le o\left(\frac{1}{1-\delta}\right)$$

as $\delta \to 1$ with $o\left(\frac{1}{1-\delta}\right)$ independent of $(\Delta_t)_{t\in\mathbb{N}}$.

**Proof:** For any $M \in \mathbb{N}$,

$$\begin{aligned}
\sum_{t=T}^{\infty} \delta^{t-T}\Delta_t &= \sum_{t=T}^{T+M} \delta^{t-T}\Delta_t + \delta^{M+1} \sum_{t=T+M+1}^{\infty} \delta^{t-T-M-1}\Delta_t \\
&= \sum_{t=T}^{T+M} (\delta^{t-T} - \delta^{T+M-t})\Delta_t + \sum_{t=T}^{T+M} \delta^{T+M-t}\Delta_t + \delta^{M+1} \sum_{t=T+M+1}^{\infty} \delta^{t-T-M-1}\Delta_t \\
&= \sum_{t=T}^{T+M} (\delta^{t-T} - \delta^{T+M-t})\Delta_t + \sum_{t=1}^{T+M} \delta^{T+M-t}\Delta_t - \sum_{t=1}^{T-1} \delta^{T+M-t}\Delta_t \\
&\quad + \delta^{M+1} \sum_{t=T+M+1}^{\infty} \delta^{t-T-M-1}\Delta_t \\
&\le \underbrace{\sum_{t=T}^{T+M} (\delta^{t-T} - \delta^{T+M-t})\Delta_t}_{A} + \delta^{M+1} \sum_{t=T+M+1}^{\infty} \delta^{t-T-M-1}\Delta_t + \left(\frac{2K}{\sqrt{1-\delta}}\right) \qquad \text{(B.21)}
\end{aligned}$$

where the last line uses condition (B.20) twice. For any $M$ even, we use the following bound for $A$,

$$\begin{aligned}
\frac{|A|}{\Delta} &\le \left[ \sum_{t=T}^{T+\frac{M}{2}} \left(\delta^{t-T} - \delta^{T+M-t}\right) + \sum_{t=T+\frac{M}{2}}^{T+M} \left(\delta^{T+M-t} - \delta^{t-T}\right) \right] \\
&\le \frac{1-\delta^{\frac{M}{2}+1}}{1-\delta} - \delta^{\frac{M}{2}+1}\frac{1-\delta^{\frac{M}{2}+1}}{1-\delta} + \frac{1-\delta^{\frac{M}{2}+1}}{1-\delta} - \delta^{\frac{M}{2}+1}\frac{1-\delta^{\frac{M}{2}+1}}{1-\delta} \\
&\le 2\frac{(1-\delta^{\frac{M}{2}+1})^2}{1-\delta}.
\end{aligned}$$

Plugging this bound into inequality (B.21) and iterating forward yields

$$\sum_{t=T}^{\infty} \delta^{t-T}\Delta_t \le \Delta\frac{2(1-\delta^{\frac{M}{2}+1})^2}{(1-\delta^{M+1})(1-\delta)} + \frac{1}{1-\delta^{M+1}}\left(\frac{2K}{\sqrt{1-\delta}}\right)$$

Set $M = 2\lfloor \frac{\widehat{M}}{2} \rfloor$ where $\widehat{M}$ is the solution to $1-\delta^{\widehat{M}+1} = (1-\delta)^{\frac{1}{6}}$. Since $(1-\delta^{M+1}) \geq (1-\delta^{\frac{M}{2}+1})$ we obtain

$$\begin{aligned}\sum_{t=T}^{\infty} \delta^{t-T}\Delta_t &\leq \Delta \frac{2(1-\delta^{M+1})^2}{(1-\delta^{M+1})(1-\delta)} + \frac{1}{1-\delta^{M+1}}\left(\frac{2K}{\sqrt{1-\delta}}\right) \\ &\leq \Delta \frac{2(1-\delta)^{\frac{1}{6}}}{(1-\delta)} + \frac{1-\delta^{\widehat{M}+1}}{1-\delta^{M+1}}\left(\frac{2K}{(1-\delta)^{\frac{4}{6}}}\right) \leq o\left(\frac{1}{1-\delta}\right).\end{aligned}$$

Since the same reasoning holds replacing $\Delta_t$ by $-\Delta_t$, it follows that $\sup_{T\in\mathbb{N}} \left| \sum_{t=T}^{\infty} \delta^{t-T}\Delta_t \right| \leq o\left(\frac{1}{1-\delta}\right)$ with $o\left(\frac{1}{1-\delta}\right)$ independent of $(\Delta_t)_{t\in\mathbb{N}}$. ■

A similar result holds if we consider the maximum regret over running starting periods.

**Lemma B.5** *Fix any $\Delta \in \mathbb{R}_+$ and suppose that for all $(\Delta_t)_{t\in\mathbb{N}}$ with $\sup_{t\in\mathbb{N}} |\Delta_t| \leq \Delta$,*

$$\sup_{T\in\mathbb{N}} \left\{ \max_{T'\leq T} \left\{ \sum_{t=T'}^{T} \delta^{T-t}\Delta_t \right\} \right\} \leq \frac{K}{\sqrt{1-\delta}}$$

*for all $\delta \in (0,1)$ with constant $K$ independent of $(\Delta_t)_{t\in\mathbb{N}}$ and $\delta$. Then for all $(\Delta_t)_{t\in\mathbb{N}}$ such that $\sup_{t\in\mathbb{N}} |\Delta_t| \leq \Delta$,*

$$\sup_{T\in\mathbb{N}} \left\{ \sum_{t=T}^{\infty} \delta^{t-T}\Delta_t \right\} \leq o\left(\frac{1}{1-\delta}\right)$$

*as $\delta \to 1$ with $o\left(\frac{1}{1-\delta}\right)$ independent of $(\Delta_t)_{t\in\mathbb{N}}$.*

**Proof:** As in the proof of Lemma B.4, we have that for all $T$ and $M$,

$$\begin{aligned}\sum_{t=T}^{\infty} \delta^{t-T}\Delta_t &= \sum_{t=T}^{T+M} \delta^{t-T}\Delta_t + \delta^{M+1} \sum_{t=T+M+1}^{\infty} \delta^{t-T-M-1}\Delta_t \\ &= \sum_{t=T}^{T+M} (\delta^{t-T} - \delta^{T+M-t})\Delta_t + \sum_{t=T}^{T+M} \delta^{T+M-t}\Delta_t + \delta^{M+1} \sum_{t=T+M+1}^{\infty} \delta^{t-T-M-1}\Delta_t \\ &\leq \sum_{t=T}^{T+M} (\delta^{t-T} - \delta^{T+M-t})\Delta_t + \frac{K}{\sqrt{1-\delta}} + \delta^{M+1} \sum_{t=T+M+1}^{\infty} \delta^{t-T-M-1}\Delta_t.\end{aligned}$$

The remainder of the proof proceeds as in the proof of Lemma B.4. ■

In Section 7 we defined processes $(\lambda_{i,t}, \mu_t, \tau_{i,t})_{i\in I, t\in\mathbb{N}}$ and backward discounted regrets,

$$\begin{aligned}
\mathcal{R}^{\tau}_{i,T} &= \sum_{t=1}^{T} \delta^{T-t}(\lambda_{i,t}(Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t)) - \tau_{i,t}(\mu_t)) \\
\mathcal{R}^{I}_{T} &= \max_{T'\le T}\left\{\sum_{t=T'}^{T} \delta^{T-t}(Y_t(I) - Y_t(\mu_t))\right\} \\
\mathcal{R}_{1,i,T} &= \max_{T'\le T}\left\{-\sum_{t=T'}^{T} \lambda_{i,t}\delta^{T-t}(Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t))\right\} \\
\mathcal{R}_{2,i,T} &= \max_{T'\le T}\left\{\sum_{t=T'}^{T} (1-\lambda_{i,t})\delta^{T-t}(Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t))\right\}.
\end{aligned}$$

**Lemma B.6** *For all strategy profiles,*

$$\begin{aligned}
&\sup_{T\in\mathbb{N}} \mathcal{R}_{1,i,T} \le K\left(\frac{1}{\sqrt{1-\delta}}\right), \quad \sup_{T\in\mathbb{N}} \mathcal{R}_{2,i,T} \le K\left(\frac{1}{\sqrt{1-\delta}}\right), \\
&\sup_{T\in\mathbb{N}} |\mathcal{R}^{\tau}_{i,T}| \le K\left(\frac{1}{\sqrt{1-\delta}}\right), \quad \sup_{T\in\mathbb{N}} \mathcal{R}^{I}_{T} \le K\left(\frac{1}{\sqrt{1-\delta}}\right)
\end{aligned}$$

*for all $\delta \in (0,1)$ with constant $K$ that depends only on $|I|$ and $y_{\max}$.*

**Proof:** The reasoning is essentially the same as that presented in Proposition 4. We provide explicit steps for $\mathcal{R}_{1,i,T}$ and $\mathcal{R}_{2,i,T}$. Observe that backward discounted regrets satisfy the following recursion:

$$\begin{aligned}
\mathcal{R}_{1,i,T+1} &= \delta\mathcal{R}^{+}_{1,i,T} - \lambda_{i,T+1}\left(Y_{-i,T+1}(I\backslash i) - Y_{-i,T+1}(\mu_{T+1})\right); \\
\mathcal{R}_{2,i,T+1} &= \delta\mathcal{R}^{+}_{2,i,T} + (1-\lambda_{i,T+1})\left(Y_{-i,T+1}(I\backslash i) - Y_{-i,T+1}(\mu_{T+1})\right).
\end{aligned}$$

In addition, observe that for all values $(R,\Delta) \in \mathbb{R}^2$, $R^+ \times (\delta R^+ + \Delta - \delta R) = R^+\Delta$. Altogether, this implies the following approachability condition:

$$\begin{aligned}
&\mathcal{R}^{+}_{1,i,T} \times (\mathcal{R}_{1,i,T+1} - \delta\mathcal{R}_{1,i,T}) + \mathcal{R}^{+}_{2,i,T} \times (\mathcal{R}_{2,i,T+1} - \delta\mathcal{R}_{2,i,T}) \le \\
&\qquad \left(-\mathcal{R}^{+}_{1,i,T}\lambda_{i,T+1} + (1-\lambda_{i,T+1})\mathcal{R}^{+}_{2,i,T}\right) \times [Y_{-i,T+1}(I\backslash i) - Y_{-i,T+1}(\mu_{T+1})] = 0.
\end{aligned}$$

Let $\mathcal{R}^{\Phi}_{i,T} = (\mathcal{R}_{1,i,T}, \mathcal{R}_{2,i,T})$. We then have that

$$\left|\left|\left[\mathcal{R}^{\Phi}_{i,T+1}\right]^{+}\right|\right|^2 \le \delta^2 \left|\left|\left[\mathcal{R}^{\Phi}_{i,T}\right]^{+}\right|\right|^2 + 2\delta\left\langle \left[\mathcal{R}^{\Phi}_{i,T}\right]^{+}, \mathcal{R}^{\Phi}_{i,T+1} - \delta\mathcal{R}^{\Phi}_{i,T}\right\rangle + \left|\left|\mathcal{R}^{\Phi}_{i,T+1} - \delta\mathcal{R}^{\Phi}_{i,T}\right|\right|^2$$

$$\leq \delta \left|\left|\left[\mathcal{R}_{i,T}^{\Phi}\right]^{+}\right|\right|^{2} + 2|I|^{2} y_{\max}^{2}$$

where we used the approachability condition to go from the first to the second line. Iterating and taking supremum over $T$ yields the result. Arguments along the lines of those in Proposition 4 establish bounds for $\mathcal{R}_{i,T}^{\tau}$ and $\mathcal{R}_{T}^{I}$. ■

We can now establish Proposition 5.

**Proof of Proposition 5:** Combining Lemmas B.4, B.5 and B.6 immediately leads to (B.16)-(B.19) with $o\left(\frac{1}{1-\delta}\right)$ independent of regrets.

Fix any history $h_T$. Assume that players other than $i$ are submitting truthful reports $m_{-i} = v_{-i}$. For any messaging strategy $m_i$ of player $i$,

$$\mathbb{E}_{m_i,v_{-i}}\left[\sum_{t=T}^{\infty}\delta^{t-T}(y_{i,t}(\mu_t) - \tau_{i,t}(\mu_t))\middle| h^T\right] \leq \mathbb{E}_{m_i,v_{-i}}\left[\sum_{t=T}^{\infty}\delta^{t-T}(Y_t(\mu_t) - Y_{-i,t}(I\backslash i))\middle| h_T\right] + o\left(\frac{1}{1-\delta}\right)$$

$$\leq \mathbb{E}_{v_i,v_{-i}}\left[\sum_{t=T}^{\infty}\delta^{t-T}(Y_t(I) - Y_{-i,t}(I\backslash i))\middle| h_T\right] + o\left(\frac{1}{1-\delta}\right)$$

where the first line follows from (B.16) and (B.19), and $o\left(\frac{1}{1-\delta}\right)$ is independent of $m_i$ and $h_T$. In turn, truthtelling ensures that player $i$ can approximately guarantee this upper bound. Conditions (B.16), (B.17) and the fact that $\mathbb{E}_{v_i,v_{-i}}\left[Y_{-i,t}(I\setminus i) - Y_{-i,t}(\mu_t)\middle| h_T\right] \geq 0$ imply that,

$$\mathbb{E}_{v_i,v_{-i}}\left[\sum_{t=T}^{\infty}\delta^{t-T}(y_{i,t}(\mu_t) - \tau_{i,t}(\mu_t))\middle| h_T\right] \geq \mathbb{E}_{v_i,v_{-i}}\left[\sum_{t=T}^{\infty}\delta^{t-T}(Y_t(I) - Y_{-i,t}(I\backslash i))\middle| h_T\right] - o\left(\frac{1}{1-\delta}\right)$$

where $o\left(\frac{1}{1-\delta}\right)$ is independent of $h_T$.

The result that truthtelling is a contemporaneous perfect $\epsilon$-equilibrium follows by choosing $\delta$ close enough to 1 that $(1-\delta)o\left(\frac{1}{1-\delta}\right) < \epsilon$.

Condition (B.17) implies that under truthtelling, equilibrium allocations are approximately efficient after any history. This implies that truthtelling is contemporaneous $\epsilon$-renegotiation proof. ■

# References


Abdulkadiroglu, A. and K. Chung (2003): "Auction design with tacit collusion," .

Aoyagi, M. (2003): "Bid rotation and collusion in repeated auctions," *Journal of economic Theory*, 112, 79–105.

Athey, S. and D. Miller (2007): "Efficiency in repeated trade with hidden valuations," *Theoretical Economics*, 2, 299–354.

Athey, S. and I. Segal (2013): "An efficient dynamic mechanism," *Econometrica*, 81, 2463–2485.

Bergemann, D. and J. Välimäki (2010): "The dynamic pivot mechanism," *Econometrica*, 78, 771–789.

Blackwell, D. (1956): "An Analog of the Minimax Theorem for Vector Payoffs," *Pacific Journal of Mathematics*, 6, 1–8.

Casella, A. (2005): "Storable Votes," *Games and Economic Behavior*, 51, 391–419.

Cesa-Bianchi, N. and G. Lugosi (2006): *Prediction, Learning, and Games*, Cambridge University Press.

Chassang, S. (2013): "Calibrated incentive contracts," *Econometrica*, 81, 1935–1971.

Che, Y. and I. Gale (1998): "Standard auctions with financially constrained bidders," *The Review of Economic Studies*, 65, 1–21.

Clarke, E. (1971): "Multipart Pricing of Public Goods," *Public Choice*, 11, 17–33.

Escobar, J. and J. Toikka (2009): "A Folk Theorem with Markovian Private Information," *Working Paper, Massachusetts Institute of Technology.*

Foster, D. and R. Vohra (1999): "Regret in the Online Decision Problem," *Games and Economic Behavior*, 29, 7–35.

Groves, T. (1973): "Incentives in Teams," *Econometrica*, 41, 617–631.

Hannan, J. (1957): "Approximation to Bayes Risk in Repeated Play," *Contributions to the Theory of Games*, 3, 97–139.

Hauser, C. and H. Hopenhayn (2005): "Trading favors: Optimal exchange and forgiveness," .

Jackson, M. and H. Sonnenschein (2007): "Overcoming Incentive Constraints by Linking Decisions," *Econometrica*, 75, 241–257.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982): "Rational cooperation in the finitely repeated prisoners' dilemma," *Journal of Economic theory*, 27, 245–252.

Lee, J. and H. Sabourian (2011): "Efficient Repeated Implementation," *Econometrica*, 79, 1967–1994.

Mailath, G., A. Postlewaite, and L. Samuelson (2005): "Contemporaneous perfect epsilon-equilibria," *Games and Economic Behavior*, 53, 126–140.

Moulin, H. (1986): "Characterizations of the pivotal mechanism," *Journal of Public Economics*, 31, 53–78.

Olszewski, W. and M. Safronov (2018a): "Efficient chip strategies in repeated games," *Theoretical Economics*, 13, 951–977.

——— (2018b): "Efficient cooperation by exchanging favors," *Theoretical Economics*, 13, 1191–1231.

Radner, R. (1980): "Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives," *Journal of Economic Theory*, 22, 136–154.

Schlag, K. H. and A. Zapechelnyuk (2017): "Dynamic benchmark targeting," *Journal of Economic Theory*, 169, 145–169.

Skrzypacz, A. and H. Hopenhayn (2004): "Tacit collusion in repeated auctions," *Journal of Economic Theory*, 114, 153–169.

Vickrey, W. (1961): "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of finance*, 16, 8–37.