Prior-Free Dynamic Allocation Under Limited Liability

Sylvain Chassang Samuel Kapon*†

Princeton University New York University

September 13, 2020.

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 VCG using limited liability detail-free mechanisms that selectively ignore reports from participants who cannot make their promised payments. We emphasize the use of prior-free online optimization techniques to approximate aggregate incentive properties of VCG.

Keywords: dynamic allocation, limited liability, renegotiation proofness, online optimization.

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

†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 outputs to 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 VCG as the agents' horizon grows large. The key observation is that by selectively ignoring the reported preferences of agents who fail to make their

VCG 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 ϵ-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 ϵ-equilibrium (Radner, 1980).[1] 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 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 sampled with positive probability and 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

[1] In a model with discounting, a similar mechanism implements efficient allocation in contemporaneous perfect ϵ-equilibrium in the sense of Mailath et al. (2005).

in environments without transferability.[2] 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 the past allocations. As a result, a player's externality must incorporate 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 deal with environments in which counterfactual outcomes are not observable, and we extend the analysis from undiscounted finite games to infinite horizon games with discounting.

The paper is structured as follows. Section 2 introduces the framework. Section 3 de-

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

scribes the usual VCG 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 shows how to address renegotiation-proofness. Appendices A and B deal with limited observability and discounting.

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. 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: $\tau_{i,t} \in [0, y_{i,t}(a_t)]$. 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.

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

$$\sum_{t=1}^{N} y_{i,t}(a_t) - \tau_{i,t}.\text{[3]} \tag{1}$$

The principal seeks to maximize the efficiency of allocations.

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

[3] We extend the analysis to an infinite-horizon version of the model with discounting in Appendix B.

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.

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: it is the solution to

$$\max_{a\in A} \sum_{i\in G} v_{i,t}(a).$$

Principal's information. We assume throughout the main text that the principal observes the entire profile of potential outputs y_t in period t.

This is obviously a strong assumption. Appendix A shows how to extend the analysis to cases in which the principal observes only the output profile $y_t(a_t) = (y_{i,t}(a_t))_{i\in I}$ corresponding to the actual decision a_t. The core idea, taken directly from the literature on online bandits (see Cesa-Bianchi and Lugosi, 2006), is to use appropriate experimentation to form unbiased estimates of average counterfactual payoffs.

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 $(a_t, \tau_t)_{t\in\{1,\cdots,N\}}$ of allocations and transfers adapted to the information available to the principal.

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

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

The main text of the paper considers two main solution concepts: ϵ-Nash equilibrium, and perfect ϵ-Nash equilibrium (Radner, 1980). Following the critique of Mailath et al. (2005), we study implementation in contemporaneous perfect ϵ-equilibria in Appendix B.

3 Benchmark Results

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

Since there are no intertemporal externalities the dynamic VCG mechanism is no different from the static VCG mechanism, and consists of requesting messages $m_{i,t}$ corresponding to players' values: $m_{i,t} \in V_i \equiv [0, y_{\max}]^A$. 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 VCG 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

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

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.*

As in the case of static VCG, there may exist other equilibria.[4] 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.[5] 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 comfort. When horizon N is large, collusive strategies remain a perfect ϵ-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 in such games.

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 the payoffs obtained under efficient allocation without transfers.

4 A Limited Liability Mechanism

We now return to the case in which the limited liability constraint is imposed and construct an explicit prior-free mechanism that implements efficient allocation in ϵ-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

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

[5]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.

viewing them as a regret minimization problem; third, we show that under the resulting mechanism, truthful reporting is an ϵ-Nash equilibrium inducing efficient allocation.

4.1 An explicit mechanism

Each period, agent i's message space is the set of possible value functions: $M_i = V_i$. We denote by $\widehat{h}_t = (m_1, y_1, \cdots, m_t)$ the history of reported states 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 μ_t depends only on history h_t. The principal implements allocation $a_t = a^*(m_t|I_t)$, i.e. the optimal allocation for selected 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 μ_t, and $\tau_{i,t}$ so that they replicate key properties of VCG: 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 subset of agents $G \subset I$, and distribution $\mu \in \Delta(\mathcal{P}(I))$ over the set $\mathcal{P}(I)$ of subsets of I, let

$$Y_t(G) \equiv \sum_{i \in I} y_{i,t}(a^*(m_t|G)) \qquad Y_t(\mu) \equiv \sum_{G \in \mathcal{P}(I)} \mu(G) Y_t(G)$$

$$Y_{-i,t}(G) \equiv \sum_{j \in I \setminus i} y_{j,t}(a^*(m_t|G)) \qquad Y_{-i,t}(\mu) \equiv \sum_{G \in \mathcal{P}(I)} \mu(G) Y_{-i,t}(G)$$

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

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."

$$\text{(efficient allocation)} \qquad \mathcal{R}_T^I \equiv \sum_{t=1}^{T} Y_t(I) - Y_t(\mu_t) \leq o(T); \tag{2}$$

$$\text{(incentive alignment)} \quad \forall i \in I, \quad \mathcal{R}_{i,T}^{\tau} \equiv \sum_{t=1}^{T} Y_{-i,t}(I \setminus i) - Y_{-i,t}(\mu_t) - \tau_{i,t}(\mu_t) \leq o(T). \tag{3}$$

Regrets $\mathcal{R}_T^I$ and $\mathcal{R}_{i,T}^{\tau}$ measure the extent to which these target incentive properties are satisfied.

Mechanism design via regret minimization. We begin by examining condition (3). Define expected profits $\pi_{i,t}(\mu_t) = y_{i,t}(\mu_t) - \tau_{i,t}(\mu_t)$. Regret $\mathcal{R}_{i,T}^{\tau}$ can be written as

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

Consider the vector of regrets $\mathcal{R}_T \equiv (\mathcal{R}_T^I, \mathcal{R}_{i,T}^{\tau})_{i \in I}$. Following Blackwell (1956) we choose transfers $\tau_{i,t}$ and μ_t so that the approachability condition

$$\left\langle \mathcal{R}_T^+, \mathcal{R}_{T+1} - \mathcal{R}_T \right\rangle \leq 0 \tag{4}$$

is satisfied for all realized values of outputs $(y_{i,t})_{i \in I}$. We have that:

$$\left\langle \mathcal{R}_T^+, \mathcal{R}_{T+1} - \mathcal{R}_T \right\rangle = \overbrace{\sum_{i \in I} \left[\mathcal{R}_{i,T}^{\tau}\right]^+ \left(\pi_{i,T+1}(\mu_t) - y_{i,T+1}(I \setminus i)\right)}^{B} \tag{5}$$

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

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

$$\mu_{T+1}(I \setminus i) = \frac{\left[\mathcal{R}_{i,T}^\tau\right]^+}{\left[\mathcal{R}_T^I\right]^+ + \sum_{i \in I} \left[\mathcal{R}_{i,T}^\tau\right]^+} \tag{6}$$

and $\mu_{T+1}(I) = 1 - \sum_{i \in I} \mu_{T+1}(I \setminus i)$.[6] This ensures that term C of (5) is equal to 0.

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

$$\tau_{i,T+1}(a) = \begin{cases} y_{i,T+1}(a) & \text{if} \quad \mathcal{R}_{i,T}^\tau > 0 \\ 0 & \quad\quad \text{otherwise.} \end{cases} \tag{7}$$

This ensures that term B of (5) is less than or equal to 0.

Proposition 2 *Consider the mechanism $(\mu_t, (\tau_{i,t})_{i \in I})_{t \geq 1}$ defined by (6) and (7). 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 ϵ-Nash equilibrium of the induced game Γ. Under truthful reporting, the allocation approaches efficiency as the horizon N gets large.*

Proof: We begin by showing that $||\mathcal{R}_N^+|| = o(N)$. Using the fact that process $(\mu_t, \tau_{i,t})_{i \in I, t \geq 1}$ satisfies approachability condition (4) for all T, we have that

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

It follows that $||\mathcal{R}_N^+|| = O(\sqrt{N})$. In addition, we now show that under truthful reporting by

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

all agents, the assumption of private values implies that for all $i \in I$,

$$\mathbb{E}\left[\mathcal{R}^{\tau}_{i,N}\right] \geq -o(N). \tag{8}$$

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

$$\mathbb{E}\left[[\mathcal{R}^{\tau,-}_{i,T+1}]^2\right] \leq \mathbb{E}\left[[\mathcal{R}^{\tau,-}_{i,T}]^2\right] - 2\mathbb{E}\left[\left\langle \mathcal{R}^{\tau,-}_{i,T}, \mathcal{R}^{\tau}_{i,T+1} - \mathcal{R}^{\tau}_{i,T}\right\rangle\right] + \mathbb{E}\left[[\mathcal{R}^{\tau}_{i,T+1} - \mathcal{R}^{\tau}_{i,T}]^2\right]. \tag{9}$$

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

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

This implies that $-2\mathbb{E}\left[\left\langle \mathcal{R}^{\tau,-}_{i,T}, \mathcal{R}^{\tau}_{i,T+1} - \mathcal{R}^{\tau}_{i,T}\right\rangle\right] \leq 0$, so that (9) yields

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

Jensen's inequality implies that $\mathbb{E}\left[\mathcal{R}^{\tau,-}_{i,N}\right] \leq \sqrt{\mathbb{E}\left[\left[\mathcal{R}^{\tau,-}_{i,N}\right]^2\right]} \leq |I| y_{\max}\sqrt{N+1}$, which yields (8).

We now show that for N sufficiently large, player i can benefit at most by ϵ from deviating from truthful reporting. By (3), we have that for any messaging strategy m_i,

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

In turn, using (8) and (2) it follows that under truthtelling player i can achieve a payoff

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

The result that truthful revelation is an ϵ-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 (2). ■

The mechanism described by (6) and (7) 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 (3) and (2) 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.[7] In the next section, we build on this feature to ensure that our mechanism is renegotiation-proof.

5 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, or intentional misreporting. In principle, the mechanism

[7] Appendix A shows that this construction can be extended to settings with unobservable counterfactuals by maintaining a minimal rate of experimentation.

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 regrets to zero and start taking into account the player's preferences again.

Second, truthful revelation may no longer be an ϵ-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 (2) hold starting from any period.

Drawing on Radner (1980), we formalize our criteria for dynamic consistency as follows.[8] 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

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

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

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

[8] In a model with discounting, and discount factor going to 1, a similar mechanism yields a contemporaneous perfect ϵ-equilibrium in the sense of Mailath et al. (2005). See Appendix B for further discussion.

5.1 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. For this purpose, we want a process $(\lambda_{i,t})_{i\in I,t\geq 1}$ such that for all $i \in I$,

$$\mathcal{R}_{1,i,T} \equiv \max_{T'\leq T}\left\{-\sum_{t=T'}^{T}\lambda_{i,t}[Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t)]\right\} \leq o(T) \tag{10}$$

$$\mathcal{R}_{2,i,T} \equiv \max_{T'\leq T}\left\{\sum_{t=T'}^{T}(1-\lambda_{i,t})[Y_{-i,t}(I\backslash i) - Y_{-i,t}(\mu_t)]\right\} \leq o(T). \tag{11}$$

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

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

Condition (10) ensures that measured externalities $(\Phi_{i,t})_{t\geq 1}$ do not become large and positive. Condition (11) 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

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

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

$$|\mathcal{R}^{\tau}_{i,T}| = o(T) \tag{12}$$

$$\mathcal{R}^{I}_{T} \leq o(T). \tag{13}$$

These properties differ from original requirements (3) and (2) 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 VCG hold approximately from the perspective of any history.

5.2 An Explicit Mechanism

The standard approachability argument of Blackwell (1956) implies that conditions (10) and (11) can be satisfied by setting

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

To see how (12) and (13) 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 dot-product appearing in the approachability condition can be written as

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

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

$$\left\langle \mathcal{R}_T^+, \mathcal{R}_{T+1} - \mathcal{R}_T \right\rangle \leq 0 \tag{14}$$

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

$$\mu_{T+1}(I\backslash i) = \frac{[\lambda_{i,T+1}\mathcal{R}_T^{I\backslash i}]^+}{[\mathcal{R}_T^I]^+ + \sum_{j\in I}[\lambda_{j,T+1}\mathcal{R}_{j,T}^\tau]^+} \tag{15}$$

and $\mu_{T+1}(I) = 1 - \sum_{i\in I} \mu_{T+1}(I \setminus i)$ and setting transfers

$$\tau_{i,T+1}(a_{T+1}) = \begin{cases} y_{i,T+1}(a_{T+1}) & \text{if} \quad \mathcal{R}_{i,T}^\tau > 0 \\ 0 & \quad\quad\text{otherwise.} \end{cases} \tag{16}$$

Proposition 3 *Consider the mechanism $(\mu_t, \tau_{i,t})_{i\in I,t\geq 1}$ defined by (15) and (16). For all $\epsilon > 0$, there exists N_0 such that for all $N > N_0$, truthful revelation is a perfect ϵ-equilibrium and ϵ-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 (10), (11), (12), (13) are satisfied.

Let us begin with (10) and (11). 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 that the following approachability condition holds:

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

Inequalities (10) and (11) follow directly from this approachability condition.

We now prove (12) and (13). The fact that $(\mu_{i,t})_{i\in I,t\geq 1}$ satisfies approachability condition (14) implies that $||\mathcal{R}_T^+|| = O(\sqrt{T})$. This implies (13), as well as $\mathcal{R}_{i,T}^\tau \leq O(\sqrt{T})$. To establish (12), we need to show that $\mathcal{R}_{i,T}^\tau \geq -O(\sqrt{T})$.

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 (10) to obtain a lower bound for $\Phi_{i,T} - \Phi_{i,\underline{T}-1}$. This establishes condition (12).

We now show that target properties (10), (11), (12), (13) imply that truthtelling is an ϵ-Nash equilibrium from the perspective of every history h_T. The proof is very similar to that of Proposition 2.

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

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

where the first, second and fourth steps respectively use (12) at N and T, (11), and (13).

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

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

where we used (12) at N and T, (13), and the fact that $\mathbb{E}[Y_{-i,t}(\mu_t) - Y_{-i,t}(I\backslash i)|h_t] \leq 0$. The result follows for N large enough that $o(N) \leq \epsilon N$.

Condition (13) implies ϵ-renegotiation proofness. ■

The mechanism $(\mu_t, \tau_{i,t})_{i\in I, t\geq 1}$ described above exhibits both cautiousness and forgiveness. Consider again 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 $\mathcal{R}_T^{I\backslash i}$ 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 $\mathcal{R}_T^I = \max_{T'\leq T} \sum_{t=T'}^{T} Y_t(I) - Y_t(\mu_t)$ is large and positive, past debts are sufficiently small to ignore. Player i is essentially forgiven, and reincluded in the decision-making process.[9]

6 Discussion

Summary. This paper shows how to replicate VCG mechanisms in a prior-free way in dynamic environments with limited liability. The main insight is that by selectively excluding

[9] 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

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 ϵ-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 truthfulness is a perfect ϵ-equilibrium and the corresponding allocation is approximately efficient from the perspective of any history.

We now discuss two extensions tackled in Appendices A and B, as well as limits of our analysis.

Limited observability. An important practical concern with our analysis is that we assume the principal observes counterfactual outcomes of different decisions $a \in A$. In practice, we may worry that individual outcomes are privately observed by agents. It turns out that our mechanisms extend as is provided messages include both reported values and reported outcomes.

A second concern is that counterfactual returns need not be observable by anyone. It turns out that this difficulty has already been solved by the literature on online bandits (Cesa-Bianchi and Lugosi, 2006). By maintaining low but non-zero exploration rates, it is possible to form unbiased estimates of counterfactual regrets. As we show in Appendix A, an adjustment of our mechanism using estimated regret enforces efficient allocation in ϵ-Nash equilibrium.

Discounting. A potential criticism of our notion of dynamic consistency is that towards the end of the game, the value of deviations may be quite large relative to *remaining* payoffs. It is only because those payoffs are scaled by a coefficient $1/N$ that these deviations appear small in our analysis. We show in Appendix B that this concern vanishes once we move to an infinite horizon game with discounted payoffs. Under an appropriately adjusted mechanism, the net-present value of deviations is small regardless of the history.

Common values and out-of-equilibrium payoff guarantees. 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 in principle under common values, one player may have a positive externality on others, which means that the mechanism may have positive ex post transfers to a player. Under the assumption that in aggregate, players have negative externalities on each other, a mechanism similar to ours, optimizing over the mapping from signals to allocations, implements truthful revelation and efficient allocation in ϵ-Nash equilibrium.

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. One difficulty is that out-of-equilibrium, VCG may exhibit common value aspects: one player may have a positive externality on others, if they are systematically misreporting their value. Under strong assumptions ensuring negative externalities even out-of-equilibrium, it is possible to establish meaningful payoff guarantees. However, strong assumptions make this result unsatisfactory. Providing satisfactory payoff-guarantees off of the equilibrium path remains an open challenge.

Appendix – Intended for Online Publication

A Limited Observability

This section relaxes observability assumptions made in the main text. We first consider the case where potential outputs $y_{i,t} = (y_{i,t}(a))_{a\in A}$ are observed, but only by player i. Second, we show how to deal with the case of unobserved counterfactuals by using randomized experimentation.

A.1 Strategic Feedback

We assume that outputs $(y_{i,t}(a))_{a\in A}$ are not publicly observed, but rather are privately observed by agent i. 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 $Y_{-i,t}$.

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.

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

$$\text{(efficient allocation)} \quad \forall T, \quad \overline{\mathcal{R}}^{I}_{i,T} = \sum_{t=1}^{T} \overline{Y}_t(I) - \overline{Y}_t(\mu_t) \leq o(T) \tag{A.2}$$

Define $\overline{\pi}_{i,t}(\mu) \equiv \overline{y}_{i,t}(\mu) - \overline{\tau}_{i,t}$. 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 μ_{T+1} be the distribution over $\{I, I \setminus i \mid i \in I\}$ such that

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

Transfers are defined by

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

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 (A.1) and (A.2) hold.

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

Proof: Let $m^*_i \equiv (v_i, y_i)$ denote the truthtelling strategy. Assume that $m_{-i} = m^*_{-i}$.

By (A.1), we have that for any messaging strategy m_i,

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

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 (A.2) imply that under truthtelling player i can achieve a payoff

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

This implies that truthtelling is an ϵ-best response for N large enough. ■

A.2 Unobserved Counterfactuals

We now assume that only outcomes for the decision a_t that is actually taken in period t are realized. The literature on online bandits (see Cesa-Bianchi and Lugosi, 2006, for an overview) shows that it is still possible to minimize regrets with large probability by estimating regrets through experimentation. We outline an extension of our approach using such a strategy.[10]

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 μ_{T+1} over decision groups in $\{I, I \setminus i \,|\, i \in I\}$ take the form

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

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

where $\nu_t = t^{-a}$, with $a \in \left(0, \frac{1}{2}\right)$, μ_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.1 *For all $\eta > 0$, there exists N_0 such that for all $T > N_0$ and for all messaging strategies $m = (m_{i,t})_{i \in I, t \geq 1}$*

$$\mathbb{E}\mathcal{R}^{\tau}_{i,T} \leq \eta T \tag{A.3}$$

$$\mathbb{E}\mathcal{R}^{I}_{T} \leq \eta T. \tag{A.4}$$

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

$$\mathbb{E}||\mathcal{R}^+_{T+1}||^2 \leq \mathbb{E}||\mathcal{R}^+_{T}||^2 + 2\mathbb{E}\left\langle \mathcal{R}^+_T, \mathcal{R}_{T+1} - \mathcal{R}_T \right\rangle + \mathbb{E}||\mathcal{R}_{T+1} - \mathcal{R}_T||^2. \tag{A.5}$$

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

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

where we used the fact that $\mathbb{E}_{\mu^*_{T+1}} \left\langle \widehat{\mathcal{R}}^+_T, \widehat{\mathcal{R}}_{T+1} - \widehat{\mathcal{R}}_T \right\rangle = 0.$[11] 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 Δ_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

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

This implies that

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

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.5) yields that there exists a constant K_1 such that

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

[11]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 ϵ-Nash equilibrium of Γ.*

The proof is identical to that of Proposition 2.

B Discounting

We now show how to adapt our approach to environments with an infinite horizon and discounting. Limit results for T sufficiently large are replaced by results for δ sufficiently near 1. Players now evaluate messaging strategies according to

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

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 ϵ-equilibrium (Mailath et al., 2005), rather than perfect ϵ-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 B.1 (Mailath et al. (2005)) *A* contemporaneous perfect ϵ-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 $\tilde{m}_i$.

Under this equilibrium concept, continuation payoffs at T are evaluated from the perspective of time T, rather than time 1. A simple adjustment to the mechanism of Section 5 achieves this property.

Target properties. As in Section 5, we consider processes $(\lambda_{i,t}, \mu_t, \tau_{i,t})_{i\in I, t\in\mathbb{N}}$. For all $i \in I$, $T \in \mathbb{N}$, we define the following present values of future flow regrets:

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

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

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

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

Our goal is to show that for some choice of $(\lambda_{i,t}, \mu_t, \tau_{i,t})_{i\in I, t\in\mathbb{N}}$, we have that for all $T \in \mathbb{N}$, and $i \in I$

$$|\mathcal{P}^\tau_{i,T}| \leq o\left(\frac{1}{1-\delta}\right)$$

and

$$\left[\mathcal{P}^I_T\right]^+, \left[\mathcal{P}_{1,i,T}\right]^+, \left[\mathcal{P}_{2,i,T}\right]^+ \leq o\left(\frac{1}{1-\delta}\right)$$

The difficulty is that these present values are forward looking, and involve flow 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 bounded sequence of flow regrets in $\mathbb{R}^{\mathbb{N}}$. We define the backwards discounted cumulative regret by

$$\mathcal{R}_T = \sum_{t=1}^{T} \delta^{T-t}\Delta_t$$

Lemma B.2 *Let $(\Delta_t)_{t\in\mathbb{N}}$ be such that for all $t \in \mathbb{N}$, $|\Delta_t| \leq \Delta$ for some fixed $\Delta \in \mathbb{R}$. If as δ gets close to 1,*

$$|\mathcal{R}_T| = O\left(\frac{1}{\sqrt{1-\delta}}\right) \tag{B.6}$$

then

$$\left|\sum_{t=T}^{\infty}\delta^{t-T}\Delta_t\right| = O\left(\frac{1}{(1-\delta)^{\frac{5}{6}}}\right).$$

Proof: We have that

$$\begin{aligned}
\sum_{t=T}^{\infty}\delta^{t-T}\Delta_t &= \sum_{t=T}^{T+M}\delta^{t-T}\Delta_t + \delta^{M+1}\sum_{t=T+M+1}^{\infty}\delta^{t-T-M-1}\Delta_t \\
&= \sum_{t=T}^{T+M}(\delta^{t-T}-\delta^{T+M-t})\Delta_t + \sum_{t=T}^{T+M}\delta^{T+M-t}\Delta_t + \delta^{M+1}\sum_{t=T+M+1}^{\infty}\delta^{t-T-M-1}\Delta_t \\
&= \sum_{t=T}^{T+M}(\delta^{t-T}-\delta^{T+M-t})\Delta_t + \sum_{t=1}^{T+M}\delta^{T+M-t}\Delta_t - \sum_{t=1}^{T-1}\delta^{T+M-t}\Delta_t \\
&\quad + \delta^{M+1}\sum_{t=T+M+1}^{\infty}\delta^{t-T-M-1}\Delta_t \\
&\leq \underbrace{\sum_{t=T}^{T+M}(\delta^{t-T}-\delta^{T+M-t})\Delta_t}_{A} + \delta^{M+1}\sum_{t=T+M+1}^{\infty}\delta^{t-T-M-1}\Delta_t + O\left(\frac{1}{\sqrt{1-\delta}}\right) \qquad \text{(B.7)}
\end{aligned}$$

where the last line uses condition (B.6) twice. We use the following bound for A:

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

Plugging this bound into inequality (B.7) and iterating forward yields.

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

Set $M+1 = \frac{\log(1-(1-\delta)^{\frac{1}{6}})}{log(\delta)}$ (which implies that $1-\delta^{M+1} = (1-\delta)^{\frac{1}{6}}$). Observing that

$$(1-\delta^{M+1}) \geq (1-\delta^{\frac{M}{2}+1})$$

we obtain

$$\sum_{t=T}^{\infty}\delta^{t-T}\Delta_t \le \Delta\frac{(1-\delta^{\frac{M}{2}+1})^2}{(1-\delta^{\frac{M}{2}+1})(1-\delta)} + O\left(\frac{1}{(1-\delta)^{\frac{5}{6}}}\right)$$
$$\le \Delta\frac{(1-\delta)^{\frac{1}{6}}}{(1-\delta)} + O\left(\frac{1}{(1-\delta)^{\frac{5}{6}}}\right) \le O\left(\frac{1}{(1-\delta)^{\frac{5}{6}}}\right)$$

Since the same reasoning holds replacing Δ_t by $-\Delta_t$, it follows that $\sum_{t=T}^{\infty}\delta^{t-T}\Delta_t \ge -O\left(\frac{1}{(1-\delta)^{\frac{5}{6}}}\right)$. ■

A similar result holds even if we consider the maximum regret over running starting periods. Let $\mathcal{R}_T^M \equiv \max_{T' \le T}\left\{\sum_{t=T'}^{T}\delta^{T-t}\Delta_t\right\}$.

Lemma B.3 *Take $\Delta > 0$ fixed, and assume that there exists K fixed such that for all sequences $(\Delta_t)_{t\in\mathbb{N}}$ satisfying $|\Delta_t| \le \Delta$ for all t,*

$$\mathcal{R}_T^M \le \frac{K}{\sqrt{1-\delta}}.$$

Then there exists a constant K' such that for all T and all such sequences $(\Delta_t)_{t\in\mathbb{N}}$,

$$\sup_T\left\{\sum_{t=T}^{\infty}\delta^{t-T}\Delta_t\right\} \le \frac{K'}{1-\delta^{\frac{5}{6}}}.$$

Proof: As in the proof of Lemma B.2, we have that for all T,

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

The remainder of the proof proceeds as in the proof of Lemma B.2. ■

A mechanism. Given processes $(\lambda_{i,t}, \mu_t, \tau_{i,t})_{i\in I, t\in\mathbb{N}}$, define backward discounted regrets

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

Processes $(\lambda_{i,t}, \mu_t, \tau_{i,t})_{i\in I, t\in\mathbb{N}}$ are defined as in Section 5, replacing undiscounted regrets with the backward discounted regrets defined above.

Lemma B.4 *We have that:*

$$\mathcal{R}_{1,i,T} \leq O\left(\frac{1}{\sqrt{1-\delta}}\right); \qquad \mathcal{R}_{2,i,T} \leq O\left(\frac{1}{\sqrt{1-\delta}}\right);$$

$$|\mathcal{R}^\tau_{i,T}| = O\left(\frac{1}{\sqrt{1-\delta}}\right); \qquad \mathcal{R}^I_T \leq O\left(\frac{1}{\sqrt{1-\delta}}\right).$$

Proof: The reasoning is essentially the same as the one presented in Section 5. 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:

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

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

$$\begin{aligned}
\left|\left|\left[\mathcal{R}^\Phi_{i,T+1}\right]^+\right|\right|^2 &\leq \delta^2\left|\left|\left[\mathcal{R}^\Phi_T\right]^+\right|\right|^2 + \delta\left\langle\left[\mathcal{R}^\Phi_T\right]^+, \mathcal{R}^\Phi_{i,T+1} - \delta\left[\mathcal{R}^\Phi_T\right]^+\right\rangle + 2I^2y^2_{\max} \\
&\leq \delta\left|\left|\left[\mathcal{R}^\Phi_T\right]^+\right|\right|^2 + 2I^2y^2_{\max}.
\end{aligned}$$

Iterating, this implies that $\mathcal{R}^\Phi_T \leq O\left(1/\sqrt{1-\delta}\right)$. Similar arguments establish bounds for $\mathcal{R}^\tau_{i,T}$, and $\mathcal{R}^I_T$. ■

An immediate corollary of this result and Lemmas B.2 and B.3 is that

$$|\mathcal{P}^{\tau}_{i,T}| \leq O\left(\frac{1}{(1-\delta)^{\frac{5}{6}}}\right) \tag{B.8}$$

$$\left[\mathcal{P}^{I}_{T}\right]^{+} \leq O\left(\frac{1}{(1-\delta)^{\frac{5}{6}}}\right) \tag{B.9}$$

$$[\mathcal{P}_{1,i,T}]^{+} \leq O\left(\frac{1}{(1-\delta)^{\frac{5}{6}}}\right) \tag{B.10}$$

$$[\mathcal{P}_{2,i,T}]^{+} \leq O\left(\frac{1}{(1-\delta)^{\frac{5}{6}}}\right) \tag{B.11}$$

We can now state the main result of this Section. Recall that $\Gamma(\delta)$ is the game induced by this mechanism.

Proposition B.3 *For all $\epsilon > 0$, there exists $\underline{\delta}$ such that for all $\delta \in (\underline{\delta}, 1)$, truthful revelation is a contemporaneous perfect ϵ-equilibrium of game $\Gamma(\delta)$.*

Proof: 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,

$$\begin{aligned}\mathbb{E}_{m_i,v_{-i}}\left[\sum_{t=T}^{\infty}\delta^{t-T}(y_{i,t}(\mu_t)-\tau_{i,t}(\mu_t))|h^T\right] &\leq \mathbb{E}_{m_i,v_{-i}}\left[\sum_{t=T}^{\infty}\delta^{t-T}(Y_t(\mu_t)-Y_{-i,t}(I\backslash i))|h_T\right]+O\left(1/(1-\delta)^{\frac{5}{6}}\right)\\ &\leq \mathbb{E}_{v_i,v_{-i}}\left[\sum_{t=T}^{\infty}\delta^{t-T}(Y_t(I)-Y_{-i,t}(I\backslash i))|h_T\right]+O\left(1/(1-\delta)^{\frac{5}{6}}\right)\end{aligned}$$

where the first line follows from (B.8) and (B.11) and the second line follows from the fact $\mathbb{E}_v\left[Y_t(I)-Y_t(\mu_t)\right] \geq 0$. In turn, truthtelling ensures that player i can approximately guarantee this upper bound. Conditions (B.8), (B.9) and the fact that $\mathbb{E}\left[Y_{-i,t}(I \setminus i) - Y_{-i,t}(\mu_t)\right] \geq 0$ imply that

$$\mathbb{E}_{v_i,v_{-i}}\left[\sum_{t=T}^{\infty}\delta^{t-T}(y_{i,t}(\mu_t)-\tau_{i,t}(\mu_t))\Big|h_T\right] \geq \mathbb{E}_{v_i,v_{-i}}\left[\sum_{t=T}^{\infty}\delta^{t-T}(Y_t(I)-Y_{-i,t}(I\backslash i))\Big|h_T\right]-O\left(1/(1-\delta)^{\frac{5}{6}}\right)$$

This implies that from the perspective of any history h_T, the payoff from truth-telling is within $O\left(1/(1-\delta)^{\frac{5}{6}}\right)$ of the payoff from any messaging strategy (uniformly over histories). Since $(1-\delta)/(1-\delta)^{5/6}$ approaches 0 as δ approaches 1, this implies that truthtelling is a contemporaneous perfect ϵ-equilibrium. ■

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.

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.