Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

On the Futility of Dynamics in Robust Mechanism Design

Santiago R. Balseiro, Anthony Kim, Daniel Russo

To cite this article:
Santiago R. Balseiro, Anthony Kim, Daniel Russo (2021) On the Futility of Dynamics in Robust Mechanism Design. Operations Research

Published online in Articles in Advance 01 Oct 2021

. https://doi.org/10.1287/opre.2021.2122 http://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Articles in Advance, pp. 1–17
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Crosscutting Areas

On the Futility of Dynamics in Robust Mechanism Design

Santiago R. Balseiro,[a] Anthony Kim,[b] Daniel Russo[a]

[a] Columbia University, New York, New York 10027; [b] Amazon, New York, New York 10001

Contact: srb2155@columbia.edu, https://orcid.org/0000-0002-0012-3292 (SRB); tonyekim@gmail.com, https://orcid.org/0000-0001-8908-2660 (AK); djr2174@gsb.columbia.edu, https://orcid.org/0000-0001-5926-8624 (DR)

Received: November 12, 2019
Revised: October 5, 2020
Accepted: December 21, 2020
Published Online in Articles in Advance: October 1, 2021

Subject Classifications: stochastic; games/group decisions
Area of Review: Revenue Management and Market Analytics

https://doi.org/10.1287/opre.2021.2122

Abstract. We consider a principal who repeatedly interacts with a strategic agent holding private information. In each round, the agent observes an idiosyncratic shock drawn independently and identically from a distribution known to the agent but not to the principal. The utilities of the principal and the agent are determined by the values of the shock and outcomes that are chosen by the principal based on reports made by the agent. When the principal commits to a dynamic mechanism, the agent best-responds to maximize his aggregate utility over the whole time horizon. The principal's goal is to design a dynamic mechanism to minimize his worst-case regret, that is, the largest difference possible between the aggregate utility he could obtain if he knew the agent's distribution and the actual aggregate utility he obtains. We identify a broad class of games in which the principal's optimal mechanism is static without any meaningful dynamics. The optimal dynamic mechanism, if it exists, simply repeats an optimal mechanism for a single-round problem in each round. The minimax regret is the number of rounds times the minimax regret in the single-round problem. The class of games includes repeated selling of identical copies of a single good or multiple goods, repeated principal-agent relationships with hidden information, and repeated allocation of a resource without money. Outside this class of games, we construct examples in which a dynamic mechanism provably outperforms any static mechanism.

Funding: The work of A. Kim was supported in part by a Columbia DRO Postdoctoral Award and was done prior to joining Amazon.
Supplemental Material: The online appendices are available at https://doi.org/10.1287/opre.2021.2122.

Keywords: strategic learning • robust mechanism design • minimax regret • dynamic pricing • dynamic contracting

1. Introduction

Individuals increasingly have repeated interactions with the same online platform. Commuters check the same ride-hailing app each morning, freelancers frequently hunt for short-term work on the same online marketplace, and advertisers bid daily on the same ad exchange. These interactions generate data the platform could use to personalize future offerings. Sometimes, the objectives of the platform and users are aligned—such as when an online music service recommends songs tailored to a user's tastes and the user's experience improves as accurate data are gathered. But often, the incentives of the platform and users are misaligned and repeated interactions become more complicated when a user strategically responds to the platform's strategy. Consider a platform that targets discount coupons at users who appear price-sensitive. This incentivizes loyal price-insensitive customers to mimic those who are not, complicating any inference from past data. Similar concerns arise in online ad exchanges—where, due to ad targeting, a meaningful fraction of auctions contain only a single bidder with a significantly high bid and appropriately setting reserve prices is a key driver of revenue—or online freelancing platforms—where a freelancer might reject an otherwise profitable contract to avoid signaling they are open to working for a low wage in the future.

In such environments, the platform could employ a myriad of dynamic strategies under which the offers available to an individual depend on all the past interactions. How much additional benefit can be derived from such dynamic strategies when an individual is strategic? We use the language of robust mechanism design to formalize a stark impossibility result. We identify a broad class of problems in which an optimal dynamic mechanism is static and simply repeats a single-round mechanism over and over. In this sense, the platform cannot benefit by using a more complex mechanism with meaningful dynamics, including any schemes that attempt to infer the private information of an individual and exploit this information using, for example, dynamic schemes (Bakos and Brynjolfsson 1999, Jackson and Sonnenschein 2007) that link together outcomes across periods. Intuitively, dynamic mechanisms that adapt based on previous actions can be manipulated by a strategic individual to induce future outcomes that are beneficial for him at the expense of the platform. Therefore, the platform finds it optimal to commit to implementing a static mechanism that does not exploit the individual's private information beyond what is known at the beginning of their repeated interactions.

For these problems, our results could be interpreted negatively as showing the impossibility of learning and exploiting the private information of a strategic individual. Viewed more positively, these results lead to massive simplification in that static mechanisms are not only robust to strategic manipulations but also optimal, allowing the platform to search over the more tractable space of single-round mechanisms. In addition, these results justify the use of simple mechanisms with substantial practical advantages: static mechanisms are simple to implement and alleviate the need for individuals to engage in complex strategic behavior. Interestingly, for some other problems, it is still possible for the platform to implement a dynamic mechanism and perform strictly better than implementing static mechanisms.

1.1. Contributions

We study a model where a principal and a strategic agent repeatedly play a game over a discrete-time finite horizon of length T. In each round, the agent privately observes an idiosyncratic shock drawn independently and identically from a distribution known to the agent but not to the principal, and the principal and agent interact through the game to realize an outcome and respective utilities. Both parties derive aggregate utility equal to the sum of their utilities across individual rounds. When the principal commits to a dynamic mechanism, the agent is strategic in the sense that he plays a best-response strategy to maximize his aggregate utility. Drawing inspiration from the enormous literature on dynamic learning in nonstrategic environments (see, e.g., Kleinberg and Leighton 2003, Besbes and Zeevi 2009), we measure the performance of a dynamic mechanism through its worst-case regret, that is, the largest difference possible between the aggregate utility he could obtain if he knew the agent's distribution and the actual aggregate utility he obtains. The principal's objective is minimax regret and the principal designs a dynamic mechanism to minimize the worst-case regret.

We provide false-dynamics results for a broad class of games, showing the principal's optimal mechanism is static without any meaningful dynamics. More specifically, we show the minimax regret is T times the minimax regret of a single-round problem and repeating T times a (near) optimal single-round mechanism from the single-round problem is correspondingly (near) optimal in the multiround problem. We prove our results under two assumptions. First, the set of possible distributions for the agent includes all point masses, that is, the principal must guard against the possibility that the agent's preferences are constant over time. Second, the optimal performance achievable by the principal with the knowledge of the agent's distribution should be *extreme-point convex*, that is, for any possible agent's distribution, the optimal performance achievable for that distribution is at most the convex combination of optimal performances corresponding to point masses where the convex combination is determined by the distribution.

Our analysis relies on leveraging point-mass distributions as worst-case distributions. When restricted to point-mass distributions, we obtain a static information structure where the agent's shock is constant and an optimal dynamic mechanism for the principal is static and repeats a single-round mechanism. Under the extreme-point convexity assumption, the optimality of static mechanisms over point masses as worst-case distributions extends to all possible distributions. We explain the extreme-point convexity assumption in terms of two opposing effects of shock uncertainty—information asymmetry and trade across shocks—and provide sufficient conditions for the assumption to hold. To the best of our knowledge, the second effect of trade across shocks is novel and may be of independent interest.

For specific applications of our general false-dynamics results, we consider (1) the dynamic selling mechanism design problem where a seller sells independent units of a single or multiple goods sequentially over time to a buyer and maximizes revenue or welfare, (2) the principal-agent model with hidden costs where a principal has a nonlinear revenue function and repeatedly contracts with an agent to produce at particular output levels, and (3) the repeated resource allocation problem without monetary transfers where a social planner allocates a costly resource in settings where monetary transfers are not allowed. In all these applications, our assumptions hold and an optimal mechanism for the multiperiod problem simply repeats an optimal mechanism for a single-round problem.

When our assumptions do not hold, it is possible that static mechanisms are not optimal and we show specific games in which either assumption does not hold and a dynamic mechanism provably outperforms any static mechanism. Finally, we extend our results in several directions and discuss connections to other related settings: a multiplicative performance guarantee, saddle-point properties, alternative benchmarks, serially correlated shock processes, a stronger notion of regret, and the maximin utility objective.

1.2. Related Work

We discuss connections between our work and several streams of literature including Bayesian dynamic mechanism design, robust mechanism design, and strategic learning.

1.2.1. Bayesian Mechanism Design. This stream of literature studies Bayesian mechanism design problems where the principal and agent share a common known prior over the distribution of shocks. False dynamics is a recurring phenomenon where optimal mechanisms do not display meaningful dynamics in sequential problems with static information (Laffont and Tirole 1993, Börgers et al. 2015). It was first observed by Baron and Besanko (1984) who considered a continuing relationship between a firm who reports cost information and a regulator who grants a license to operate. Similar results hold in many other dynamic allocation models where the agent's shock is constant (e.g., Baron and Besanko 1984) or changing over time (e.g., Bakos and Brynjolfsson 1999, Kakade et al. 2013, Pavan et al. 2014). Although generally dependent on the time horizon, optimal mechanisms may determine the outcomes of all future periods in the first round in these results. To the best of our knowledge, our paper is the first to study false-dynamics in a general class of problems with respect to a minimax regret objective. Although our model is more aligned with sequential screening models (e.g., Courty and Li 2000, Krähmer and Strausz 2015, Bergemann et al. 2017) where the optimal mechanisms are dynamic, we still obtain false-dynamics results. Under more stringent liquidity or participation constraints, false-dynamics disappears and adaptive mechanisms can outperform the optimal static mechanism in Bayesian settings (Krishna et al. 2013, Ashlagi et al. 2016, Balseiro et al. 2018). Our results reveal a stronger collapse of dynamics for a class of problems in that they hold even when these participation constraints are imposed.

1.2.2. Robust Mechanism Design. Although Bayesian models have appealing philosophical foundations, the resulting mechanisms sometimes place impractical requirements on the prior information of the designer. Wilson (1987) argues that mechanisms should not excessively rely on probabilistic assessments on the agents' types. Our work contributes to the robust mechanism design literature that was pioneered by Bergemann and Schlag (2008, 2011). In particular, the authors consider the single-round problem where the principal (i.e., seller) sells a good to an agent (i.e., buyer) to minimize the worst-case regret without the knowledge of the agent's distribution in Bergemann and Schlag (2008). Carrasco et al. (2019) is perhaps the most closely related work to ours. They show a similar false-dynamics result in an auction setting with respect to the maximin utility objective where the principal knows the mean of the unknown distribution and maximizes the worst-case utility over distributions that are potentially correlated over time. They bound the worst-case utility of any dynamic mechanisms by considering a worst-case distribution that is perfectly correlated across time and then invoking standard false-dynamics results from the Bayesian literature. Our approach in the minimax regret setting relies on considering point masses as worst-case distributions, which are not feasible in Carrasco et al. (2019) because of the moment constraints, and applies for a broader class of games. There are many other works in this literature (e.g., Carrasco et al. 2019; Kos and Messner 2015; Carroll 2017; Pınar and Kızılkale 2017; Carrasco et al. 2018a, b; Kocyigit et al. 2018), but they consider single-round problems whereas our problem is a multiround problem.

1.2.3. Strategic Learning. For a special case of our general problem, Amin et al. (2013) has previously shown that regret must grow linearly with the time horizon. Such results show that the principal bears a cost of asymmetric information that does not vanish regardless of the length of the time horizon. This formalizes the common folklore that learning about a strategic agent is fundamentally more difficult than learning about a myopic one. However, such results tend to be only asymptotic in nature and do not speak to the exact, absolute potential benefits of dynamic mechanisms over static ones, which is our main contribution.

Numerous papers in the learning theory literature consider assumptions that enable efficient dynamic learning on the part of the principal. Most notably, positive results are available when the principal repeatedly interacts with a myopic agent who optimizes without internalizing future consequences of his actions or a population of agents who each interact with the principal only once or when the principal simultaneously interacts with multiple agents whose values are drawn i.i.d. from the same distribution (e.g., Kleinberg and Leighton 2003, Kanoria and Nazerzadeh 2020). When the principal interacts with an agent who is forward-looking but less patient, usually modeled through unequal discount factors, the principal can learn and exploit the agent's private information to the extent that their time-preferences differ (e.g., Amin et al. 2013, 2014; Mohri and Munoz 2014, 2015; Golrezaei et al. 2020). Not surprisingly, the performance guarantees obtained in these settings degrade as the difference between the principal's and agent's discount factors becomes small. A distinguishing feature of our model is that the principal and agent are placed on a more equal footing in that they are both forward-looking and equally patient (i.e., the same discount factor). Many online platforms are characterized by short planning horizons and a high frequency of transactions. Although both parties might discount future payoffs in the one-to-one relationship, the difference in the discount factors can be expected to be small as planning horizons span only for weeks or months.

2. Model

We consider games described in terms of a time horizon T and an environment (Ω, Θ, u, v) where Ω is a set of outcomes, Θ is a set of idiosyncratic shocks of the agent, $u : \Theta \times \Omega \to \mathbb{R}$ is the utility function of the principal, and $v : \Theta \times \Omega \to \mathbb{R}$ is the utility function of the agent. A principal and an agent repeatedly interact in the given environment over T rounds, producing outcomes $\omega_1, \ldots, \omega_T \in \Omega$. Independently and identically distributed shocks $\theta_1, \ldots, \theta_T \sim F$ might influence the utility of both the principal and agent. The shock distribution $F \in \Delta(\Theta)$ is private to and learned by the agent in Round 0 before the shocks, and the shock θ_t is privately observed by the agent at the start of Round t. Assume that Ω contains a designated no-interaction outcome denoted by $\emptyset$. Both the principal and agent attain a utility of zero in any round with no interaction.

Example 1 (Selling Problem). The principal (or seller) repeatedly offers identical copies of an item to the agent (or buyer). The outcome $\omega_t = (x_t, p_t)$ realized in Round t consists of a quantity $x_t \geq 0$ of the good received by the agent and a payment $p_t \in \mathbb{R}$. The shock θ_t is the agent's willingness-to-pay or valuation for the good in Round t and F is his private valuation distribution. The agent's utility function is $v(\theta, (x, p)) = \theta x - p$ and the principal's is his revenue $u(\theta, (x, p)) = p$. The no-interaction outcome is one where $x = 0$ and $p = 0$. The private distribution F could be thought of as the agent's type and is persistent over time. Additional randomness in the shock θ_t, beyond F, represents unpredictable external factors that influence the agent's preferences in that round.

As is standard in the mechanism design literature, the principal can commit to implement a dynamic mechanism, which specifies a full protocol of interaction between the principal and agent.[1] Formally, a dynamic mechanism can be written as a tuple $A = (\{\mathcal{M}_t\}_{0:T}, \{\pi_t\}_{1:T}, \{\sigma_t\}_{0:T})$, where we use $a : b$ as shorthand for $a, \ldots, b$. For each t, the set $\mathcal{M}_t$ is the report space and defines the space of possible messages the agent can transmit to the principal in Round t. Let $\mathcal{M} = \{\mathcal{M}_t\}_{0:T}$. A decision rule π_t specifies an outcome $\omega_t = \pi_t(m_t, h_t, z_t) \in \Omega$ in Round t as a function of the report $m_t \in \mathcal{M}_t$, the history of prior interaction $h_t = (m_{0:t-1}, \omega_{1:t-1})$, and a private random variable z_t drawn independently over time from a uniform distribution over $[0, 1]$. Let $\pi = \{\pi_t\}_{1:T}$. The private random variable z_t allows outcomes to be determined on a randomized basis. Assuming it to be uniformly distributed is without loss of generality since more complex random variables can be generated by inverse transform sampling.

A dynamic mechanism A allows the principal to implement per-round decision rules that may be linked across rounds and depend on the history of past interactions. For example, in the selling problem, the principal may post a reserve price that he adjusts dynamically, bundle the current item and future items together, or provide some discount scheme that offers a future item at a low price if the current item is bought at a high price.

Following the convention in the mechanism design literature, the principal's mechanism specifies a recommended agent strategy $\sigma := \{\sigma_t\}_{0:T}$. We later constrain the choice of the recommended strategy such that the mechanism is incentive compatible. In each round t, σ_t specifies the report $m_t = \sigma_t(\theta_t, h_t^+, y_t) \in \mathcal{M}_t$ as a function of the realization of the private shock θ_t, the augmented history h_t^+ containing all information available (i.e., the public history h_t and the agent's private information) to the agent prior to Round t, and a private random variable y_t drawn independently from a uniform distribution over $[0, 1]$. The initial augmented history is $h_0^+ := F$ while $h_t^+ = (F, \theta_{1:t-1}, m_{0:t-1}, \omega_{1:t-1})$ for $t > 0$. Notice that no outcome is determined in Round 0, but the agent's initial message $m_0 = \sigma_0(F, y_0)$ could influence subsequent outcomes.

Given the principal's dynamic mechanism $A = (\mathcal{M}, \pi, \sigma)$, the agent's strategy $\tilde{\sigma} = \{\tilde{\sigma}_t\}_{0:T}$ (which may be different from σ), the distribution F, and the time horizon T, the principal and agent's total expected utilities are defined, respectively, as

$$\text{PrincipalUtility}(A, \tilde{\sigma}, F, T) := \mathbb{E}_{\pi, \tilde{\sigma}}\left[\sum_{t=1}^{T} u(\theta_t, \omega_t)\right]$$

and

$$\text{AgentUtility}(A, \tilde{\sigma}, F, T) := \mathbb{E}_{\pi, \tilde{\sigma}}\left[\sum_{t=1}^{T} v(\theta_t, \omega_t)\right].$$

The expectations above are taken over the realizations of the shocks $(\theta_1, \ldots, \theta_T)$ and the private random variables $\{z_t\}_{1:T}$ and $\{y_t\}_{0:T}$ which are omitted. The subscript indicates that the agent's messages are determined by $\tilde{\sigma}$ and the mechanism's outcomes are determined by π, meaning $m_t = \tilde{\sigma}_t(\theta_t, h_t^+, y_t)$ for $t \in \{0, \ldots, T\}$ and $\omega_t = \pi_t(m_t, h_t, z_t)$ for $t \in \{1, \ldots, T\}$. It is implicitly understood that the agent's augmented histories $h_0^+ := F$ and $h_t^+ = (F, \theta_{1:t-1}, m_{0:t-1}, \omega_{1:t-1})$ for $t \geq 1$ contain information about the same distribution F from which the shocks are drawn. We typically omit the subscripts from the above expectations, as they are clear from the context. We remark that the principal may not directly observe his utility if it depends on the agent's private shocks as, for example, in the case of welfare maximization. Figure 1 summarizes the order of events over the time horizon.

We say that dynamic mechanism $A = (\mathcal{M}, \pi, \sigma)$ is *incentive compatible* (IC) if the inequality

$$\text{AgentUtility}(A, \sigma, F, T) \geq \text{AgentUtility}(A, \tilde{\sigma}, F, T)$$

holds for every probability distribution F over Θ and every feasible agent strategy $\tilde{\sigma}$, that is, the agent is weakly better off following the principal's recommendation. We follow the convention in mechanism design of assuming the agent follows the recommended strategy if the mechanism is incentive compatible.[2] With this convention, we simplify the notation when A is IC and write $\text{PrincipalUtility}(A, F, T) := \mathbb{E}_{\pi,\sigma}\left[\sum_{t=1}^{T} u(\theta_t, \omega_t)\right]$ and $\text{AgentUtility}(A, F, T) := \mathbb{E}_{\pi,\sigma}\left[\sum_{t=1}^{T} v(\theta_t, \omega_t)\right]$ with the understanding that the omitted recommended strategy is the utility-maximizing strategy for the agent chosen by the principal.

We model an *individual rationality* (IR) constraint (equivalently, *participation* constraint) by providing a no-participation option in Round 0 and imposing IC with respect to this option. More specifically, we assume $\mathcal{M}_0$ contains a special report denoted by QUIT. If the agent reports QUIT in Round 0, he does not participate and the outcome is understood to be the no-interaction outcome over the whole time horizon. Because reporting QUIT is a feasible deviation, an incentive compatible mechanism should always provide the agent a nonnegative expected utility. This is commonly known as the ex-ante IR constraint in that the agent decides to participate or not while knowing only his distribution but not future shocks.

Let $\mathcal{A}$ denote the set of all incentive compatible dynamic mechanisms. If the agent has private distribution F, the optimal performance attainable by a principal who knows this distribution is

$$\text{OPT}(F, T) := \sup_{A \in \mathcal{A}} \text{PrincipalUtility}(A, F, T).$$

An optimal solution for this Bayesian dynamic mechanism design problem can be characterized recursively using the promised utility framework pioneered by Green (1987), Spear and Srivastava (1987), and Thomas and Worrall (1990). Alternatively, when the T is large, asymptotically optimal mechanisms can sometimes be provided (see, e.g., Fudenberg et al. 1994, Jackson and Sonnenschein 2007). The regret defined as

$$\text{Regret}(A, F, T) := \text{OPT}(F, T) - \text{PrincipalUtility}(A, F, T)$$

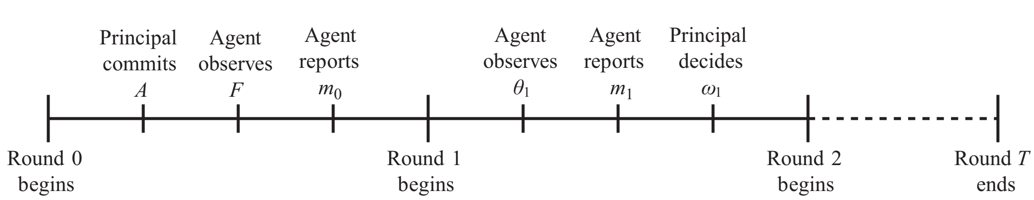
measures the shortfall in the performance of a dynamic mechanism $A \in \mathcal{A}$ against this known-distribution benchmark. The principal's objective is to design an incentive compatible mechanism with minimal worst-case regret where the worst-case regret for mechanism A is defined as

$$\text{Regret}(A, T) := \sup_{F \in \mathcal{F}} \text{Regret}(A, F, T),$$

where $\mathcal{F} \subseteq \Delta(\Theta)$ is a given set of probability distributions over Θ that are possible for the agent. We can equivalently interpret that nature is selecting a worst-case distribution F against the principal's mechanism and the principal is guarding against all such possibilities in $\mathcal{F}$. The optimal minimax regret in the multiround problem is given by $\text{Regret}(T) := \inf_{A \in \mathcal{A}} \text{Regret}(A, T)$. To ensure this quantity is well defined, we assume throughout that $\sup_{F \in \mathcal{F}} \text{OPT}(F, T) < \infty$.

Remark 1. It may seem natural to instead formulate a maximin utility problem where the principal wants to solve $\sup_{A \in \mathcal{A}} \inf_{F \in \mathcal{F}} \text{PrincipalUtility}(A, F, T)$. The challenge with this formulation is that, in some games such as dynamic selling, nature could select a distribution under which the agent does not value the good at all, leading to a maximin utility of zero. If the distribution F were known, minimizing regret is equivalent to maximizing utility, but worst-case distributions are more natural under a regret objective. Our main result also applies to other robust objectives. We can show that static mechanisms are optimal for the maximin ratio objective where the principal wants to solve $\sup_{A \in \mathcal{A}} \inf_{F \in \mathcal{F}} \text{PrincipalUtility}(A, F, T)/\text{OPT}(F, T)$. In addition, static mechanisms are optimal for a constrained maximin utility objective studied by Carrasco et al. (2019) in the dynamic selling problem. These extensions are discussed in Section 6. Our primary reason for focusing on regret instead of, say, the maximin ratio objective is that minimax regret is a widely studied objective for dynamic learning in nonstrategic environments.

Figure 1. The Order of Events over the Time Horizon



3. Optimality of Direct Static Mechanisms

In this section, we provide our main results. For a general class of games satisfying two sufficient conditions, we determine the minimax regret of the multiround problem to be T times that of a single-round minimax regret problem restricted to point-mass distributions and show an optimal dynamic mechanism is static, that is, it simply repeats a single-round direct mechanism, without any meaningful dynamics.

3.1. Direct Static Mechanisms

Our paper will focus on the special class of direct static mechanisms, denoted by $\mathcal{S}^{\times T} \subset \mathcal{A}$. These mechanisms are direct, meaning that the report spaces $\mathcal{M}_t$ contain the shock space Θ and the recommended strategy for the agent is to truthfully report his shocks to the principal. They satisfy interim IR constraints, meaning that the agent prefers to participate in each round given his information up to that point (including his shock for that round). Finally, they are simple repetitions of single-round mechanisms, meaning that the outcome in a given round is determined solely by the agent's reported shock in that round and does not depend directly on the history or the current round. To be precise, $\mathcal{S}^{\times T}$ is the space of *repeated single-round, direct, IC/IR* mechanisms, but we refer to them as *direct static mechanisms*.

More formally, an IC mechanism $A = (\mathcal{M}, \pi, \sigma)$ is an element of $\mathcal{S}^{\times T}$ if:

1. The message space allows for reporting the agent's shock or a choice to not participate in this round or, equivalently, pass: $\mathcal{M}_t = \Theta \cup \{\mathsf{PASS}\}$ for each $t \in \{1, \ldots, T\}$.
2. The mechanism honors a request to pass: $\pi_t(\mathsf{PASS}, h_t, z_t) = \emptyset$ for all h_t and z_t.
3. The recommended agent strategy is truthful reporting: $\sigma_t(\theta_t, h_t^+, y_t) = \theta_t$ holds for each possible (θ_t, h_t^+, y_t) and for $t \in \{1, \ldots, T\}$.
4. The mechanism's decision rule does not depend on the history or current round: there exists $\tilde{\pi} : \mathcal{M}_1 \times [0,1] \to \Omega$ such that $\pi_t(m_t, h_t, z_t) = \tilde{\pi}(m_t, z_t)$ for each possible (m_t, h_t, z_t) and $t \in \{1, \ldots, T\}$.

A direct static mechanism is uniquely identified by a direct mechanism for a single-round problem, that is, a problem with $T = 1$. For a single-round direct mechanism $S \in \mathcal{S}^{\times 1}$, we let $S^{\times T} \in \mathcal{S}^{\times T}$ denote the direct static mechanism that simply repeats S for T rounds. Note that rounds decouple under direct static mechanisms.

As in the case of the ex-ante IR constraint we considered previously, we implicitly enforce IR constraints by requiring the mechanism to be incentive compatible in the presence of no-participation options. Notice that for static mechanisms, we have not specified a report space $\mathcal{M}_0$ in Round 0. We can take $\mathcal{M}_0 = \{\mathsf{CONTINUE}, \mathsf{QUIT}\}$, where $\mathsf{CONTINUE}$ advances the agent to the first round and, as before, QUIT indicates a choice to not participate in any future rounds. Because the agent has the freedom to choose PASS in each round, but instead prefers to report his shock truthfully, this initial round is redundant and is included only for consistency with the general formulation.

3.2. Main Result

We prove our results for the general class of games satisfying the following two assumptions. We defer most discussion of these until Section 5. For any $\theta \in \Theta$, let δ_θ denote a point mass at θ, that is, a probability distribution with $\delta_\theta(\{\theta\}) = 1$. The first assumption states that the set of possible distributions $\mathcal{F}$ contains all point masses. Interpreted differently, it says that the principal must guard against the possibility that the agent's private shock is some unknown value in Θ and is constant across time.

Assumption 1 (Possibility of Deterministic Shocks). *For every $\theta \in \Theta$, we have $\delta_\theta \in \mathcal{F}$.*

The next assumption imposes a condition on the known-distribution benchmark $\mathrm{OPT}(F, T)$. In words, the assumption means that the principal is better off when the agent's shock is constant across time and publicly known than when shocks are random across time and are privately observed by the agent. As we discuss in Section 5, the condition holds for a broad class of games and has an interesting economic interpretation. Mathematically, it imposes the defining condition of convexity, but only with respect to the point-mass distributions—which are extreme points in the simplex of probability distributions $\Delta(\Theta)$. Note any distribution F is a convex combination of point-mass distributions where the combination is given by the distribution F, that is, F and $\mathbb{E}_{\theta\sim F}[\delta_\theta]$ are equivalent in the distributional sense.

Assumption 2 (Extreme-Point Convexity). *For all $F \in \mathcal{F}$, $\mathrm{OPT}(F, T) \le \mathbb{E}_{\theta\sim F}[\mathrm{OPT}(\delta_\theta, T)]$.*

The following theorem shows a complete reduction from the multiround problem to a single-round problem (with $T = 1$) in terms of their objective values, their (nearly) optimal solutions, and the existence of optimal solutions. Recall that $\mathrm{Regret}(T) = \inf_{A\in\mathcal{A}} \sup_{F\in\mathcal{F}} \mathrm{Regret}(A, F, T)$ is the minimax regret attainable in the multiround problem.

Theorem 1. *Suppose Assumptions 1 and 2 hold. Then,* $\mathrm{Regret}(T) = T \cdot \inf_{S\in\mathcal{S}^{\times 1}} \sup_{\theta\in\Theta} \mathrm{Regret}(S, \delta_\theta, 1)$. *Moreover, for any $\epsilon \ge 0$, if a mechanism $S \in \mathcal{S}^{\times 1}$ satisfies*

$$\sup_{\theta\in\Theta} \mathrm{Regret}(S, \delta_\theta, 1) \le \inf_{S'\in\mathcal{S}^{\times 1}} \sup_{\theta\in\Theta} \mathrm{Regret}(S', \delta_\theta, 1) + \frac{\epsilon}{T}, \quad (1)$$

then,

$$\sup_{F\in\mathcal{F}} \mathrm{Regret}(S^{\times T}, F, T) \le \mathrm{Regret}(T) + \epsilon.$$

Finally, $\arg\min_{A\in\mathcal{A}} \sup_{F\in\mathcal{F}} \mathrm{Regret}(A, F, T)$ *is empty if and only if* $\arg\min_{S\in\mathcal{S}^{\times 1}} \sup_{\theta\in\Theta} \mathrm{Regret}(S, \delta_\theta, 1)$ *is empty.*

Our result is easier to interpret when the infimum in (1) is attained, that is, there exists an optimal single-round mechanism. In this case, our main theorem says that an optimal dynamic mechanism can be constructed by first solving for an optimal direct mechanism for the single-round problem $S^* \in \arg\min_{S\in\mathcal{S}^{\times 1}} \sup_{\theta\in\Theta} \mathrm{Regret}(S, \delta_\theta, 1)$ and then simply repeating this mechanism over the rounds, that is, implementing $(S^*)^{\times T}$. Such a mechanism is myopic and nonadaptive: it plans over a single round even when there are many and it does not adjust its decision rule based on prior interactions with the same agent. This provides insight into the character of an optimal mechanism and also lets us formulate more tractable optimization problems. Additionally, our result states that the minimax regret in the multiround problem is T times the minimax regret in the single-round problem, that is, $\mathrm{Regret}(T) = T\cdot\mathrm{Regret}(1)$. The final part of the result concerns the existence of direct static mechanisms that are exactly optimal.

Depending on the value of the minimax regret for the single-round problem, different interpretations are possible. If $\mathrm{Regret}(1) = 0$, repeating a single-round mechanism obtains essentially the optimal performance achievable with the knowledge of the agent's private distribution. Then, the distributional information is not necessary and learning/adaptive schemes are not beneficial to begin with and the multiround problem is easy. On the other hand, if $\mathrm{Regret}(1) > 0$, the minimax regret for the multiround problem is linear in the time horizon and repeating a single-round mechanism is still a (near) optimal dynamic mechanism. Even though knowing the agent's private distribution would be valuable, it is impossible to increase the principal's utility by employing an adaptive learning scheme. Welfare maximization in the dynamic selling mechanism design problem (Section 4.1) is of the former kind. Revenue maximization in the same problem, the principal-agent contract model (Section 4.2) and the dynamic resource allocation problem without monetary transfers (Section TR.5 of the technical report Balseiro et al. 2019) are of the latter. We defer further details to respective sections.

3.3. Characterizing Direct Static Mechanisms

Theorem 1 reduces a complex dynamic mechanism design problem to the following single-round optimization problem over direct mechanisms and restricted to point-mass distributions:

$$\inf_{S\in\mathcal{S}^{\times 1}} \sup_{\theta\in\Theta} \mathrm{Regret}(S, \delta_\theta, 1). \tag{2}$$

In this subsection, we describe the structure of (2) in more detail.[3] We first describe the space of single-round direct mechanisms in a more explicit form. Recall from Subsection 3.1 that $S\in\mathcal{S}^{\times 1}$ is uniquely determined by decision rule $\pi_1 : \Theta\cup\{\mathsf{PASS}\}\times[0,1]\to\Omega$ that determines the outcome $\omega_1 = \pi_1(\theta_1, z_1)$ as a function of the reported shock θ_1 and the random variable z_1 that enables randomized decision rules. Put differently, S is effectively defined by a rule that associates each possible report $\theta\in\Theta$ with a distribution over outcomes; we ignore the possibility of a report of PASS under the recommended truthful reporting strategy. We make this explicit and, abusing notations, equivalently define a single-round direct mechanism $S\in\Delta(\Omega)^\Theta$ by

$$S_\theta(W) = \mathbb{P}_{z\sim\mathrm{Uniform}[0,1]}(\pi_1(\theta, z)\in W),$$

for every shock $\theta\in\Theta$ and measurable set $W\subseteq\Omega$.

Consider the following optimization problem in terms of the outcome distribution representation:

$$\inf_{S\in\Delta(\Omega)^\Theta} \sup_{\theta\in\Theta} \left\{\mathrm{OPT}(\delta_\theta, 1) - \int_\Omega u(\theta,\omega)\mathrm{d}S_\theta(\omega)\right\} \tag{3}$$

$$\text{s.t.} \int_\Omega v(\theta,\omega)\mathrm{d}S_\theta(\omega) \ge \int_\Omega v(\theta,\omega)\mathrm{d}S_{\theta'}(\omega)\ \forall\ \theta,\theta'\in\Theta, \tag{IC}$$

$$\int_\Omega v(\theta,\omega)\mathrm{d}S_\theta(\omega) \ge 0 \qquad \forall\ \theta\in\Theta. \tag{IR}$$

It is simple to conclude that the constraints (IC) and (IR) are a rewriting of the incentive compatibility and (implicit) individual rationality constraints we have placed on the set $\mathcal{S}^{\times 1}$ in Subsection 3.1. The following lemma leverages this representation of single-round direct mechanisms to simplify (2) and shows its equivalence to (3); see Online Appendix A.2 for the proof.

Lemma 1. *The optimization problems* (2) *and* (3) *attain the same objective value. Moreover, a single-round direct mechanism* S^* *with decision rule* $\pi_1 : \Theta\cup\{\mathsf{PASS}\}\times[0,1]\to\Omega$ *is an optimal solution of* (2) *if and only if its outcome distributions* $S^*_\theta(W) := \mathbb{P}_{z\sim\mathrm{Uniform}[0,1]}(\pi_1(\theta,z)\in W)$ *for* $\theta\in\Theta, W\subseteq\Omega$ *are an optimal solution of* (3). *Finally, the objective of* (3) *can be equivalently replaced with* $\sup_{F\in\Delta(\Theta)}\{\int_\Theta \mathrm{OPT}(\delta_\theta,1)dF(\theta) - \int_\Theta\int_\Omega u(\theta,\omega)dS_\theta(\omega)dF(\theta)\}$.

It is easy to see that the known-distribution benchmark in (3) can be equivalently written as

$$\mathrm{OPT}(\delta_\theta, 1) = \sup_{G\in\Delta(\Omega)} \int_\Omega u(\theta,\omega)\mathrm{d}G(\omega)$$
$$\text{s.t.} \int_\Omega v(\theta,\omega)\mathrm{d}G(\omega) \ge 0.$$

Therefore, $\mathrm{OPT}(\delta_\theta, 1)$ can be thought of as a "first-best" benchmark without IC constraints in which the principal chooses, for the given shock, the best possible distribution over outcomes subject to an ex-ante IR constraint.

When the spaces of shocks and of outcomes are discrete, (3) is a finitely-sized linear program that can be efficiently solved. When the shock space is single-dimensional, an optimal single-round direct mechanism can be sometimes determined analytically using the Myersonian theory (which involves using the envelope theorem to simplify the IC constraints) and minimax duality theory (on the version of (3) with the alternative objective stated in Lemma 1 in which the inner supremum is a convex program). We illustrate this approach in Section 4. We are not aware of a general approach to solve (3) in general, multidimensional mechanism design problems.

3.4. Proof Sketch for Theorem 1

To prove our main theorem, we show a lower bound and an upper bound on the minimax regret of the multiround problem in terms of direct static mechanisms. Essentially, the multiround problem reduces to a static problem when the principal restricts to point-mass distributions that are possible candidates for the agent's distribution under Assumption 1 and this restriction is without loss for the principal under Assumption 2. Both the lower bound and upper bound arguments rely crucially on how the single-round benchmark $\mathrm{OPT}(\delta_\theta, 1)$ relates to the optimal performance achievable $\mathrm{OPT}(F, T)$ in T rounds. As already mentioned, point-mass distributions will be important and we have the following result on the benchmark. When the distribution F is a point-mass and is known to the principal, the agent holds no private information and the principal can (nearly) attain $\mathrm{OPT}(\delta_\theta, T)$ by simply repeating a (near) optimal mechanism that (nearly) attains $\mathrm{OPT}(\delta_\theta, 1)$ over T rounds.

Proposition 1. *For any* $\theta \in \Theta$, $\mathrm{OPT}(\delta_\theta, T) = T \cdot \mathrm{OPT}(\delta_\theta, 1)$.

The following result lower bounds the regret of any incentive compatible dynamic mechanism in terms of the regret of direct static mechanisms restricted to point-mass distributions and it directly implies a lower bound on the minimax regret in the multiround problem.

Lemma 2 (Lower Bound). *Suppose Assumption 1 holds. For any incentive compatible dynamic mechanism* $A \in \mathcal{A}$*, there exists a single-round direct IC/IR mechanism* $S \in \mathcal{S}^{\times 1}$ *such that*

$$\sup_{F \in \mathcal{F}} \mathrm{Regret}(A, F, T) \geq T \cdot \sup_{\theta \in \Theta} \mathrm{Regret}(S, \delta_\theta, 1). \tag{4}$$

To see the lower bound $\mathrm{Regret}(T) \geq T \cdot \inf_{S \in \mathcal{S}^{\times 1}} \sup_{\theta \in \Theta} \mathrm{Regret}(S, \delta_\theta, 1)$, first take the infimum over all single-round direct IC/IR mechanisms S on the right-hand side and then take the infimum over all incentive compatible dynamic mechanisms A on the left-hand side. To prove the lemma, we use a revelation-principle-type argument to reduce the multiround problem to the single-round problem. The main idea involves using Assumption 1 to focus on point-mass distributions and then imposing structural constraints (the IC/IR constraints) as we effectively shrink the time horizon; this idea also appears in Amin et al. (2013). More specifically, we can construct a single-round direct mechanism S from any incentive compatible dynamic mechanism A by letting S_θ for $\theta \in \Theta$ to be the time-averaged distribution of outcomes when the agent's distribution is the point-mass distribution δ_θ and the agent plays the recommended strategy (as given in A). By construction, truthfully reporting a shock θ under S gives the same utilities to both parties, when scaled by T, as implementing the recommended strategy under A. Then, the resulting single-round mechanism S is incentive compatible and individually rational because the recommended strategy is utility-maximizing for the agent and guarantees the agent utility of at least 0 for the point-mass distributions in the multiround problem.

The next result upper bounds the regret of direct static mechanisms in the multiround problem under Assumption 2. For any single-round direct IC/IR mechanism, the regret incurred by repeating it T times is no greater than T times its regret in the single-round problem when restricted to point-mass distributions.

Lemma 3 (Upper Bound). *Suppose Assumption 2 holds. For every single-round direct IC/IR mechanism* $S \in \mathcal{S}^{\times 1}$,

$$\sup_{F \in \mathcal{F}} \mathrm{Regret}\left(S^{\times T}, F, T\right) \leq T \cdot \sup_{\theta \in \Theta} \mathrm{Regret}(S, \delta_\theta, 1). \tag{5}$$

The lemma implies the upper bound $\mathrm{Regret}(T) \leq T \cdot \inf_{S \in \mathcal{S}^{\times 1}} \sup_{\theta \in \Theta} \mathrm{Regret}(S, \delta_\theta, 1)$ because we can take the infimum over all single-round direct IC/IR mechanisms on both sides in the stated inequality and note repetitions of single-round direct IC/IR mechanisms are a subset of all incentive compatible dynamic mechanisms $\mathcal{A}$. For a sketch of the proof of the lemma, we note that when the principal implements $S^{\times T}$ for a single-round direct IC/IR mechanism S and the agent reports truthfully (as recommended for direct mechanisms), the individual rounds correspondingly decouple and $\mathrm{PrincipalUtility}(S^{\times T}, F, T) = T \cdot \mathrm{PrincipalUtility}(S, F, 1)$ for any distribution F. We then note Assumption 2 guarantees that the worst-case regret against point-mass distributions extends to that against any distributions in $\mathcal{F}$ in the single-round problem. That is, by the extreme-point convexity assumption, the principal can control his regret by protecting against all point-mass distributions.

Combining Lemmas 2 and 3, we can prove Theorem 1. From the above discussion, we already have $\mathrm{Regret}(T) = T \cdot \inf_{S \in \mathcal{S}^{\times 1}} \sup_{\theta \in \Theta} \mathrm{Regret}(S, \delta_\theta, 1)$. Similarly, we can prove the second part about the (near) optimality of direct static mechanisms and the third part about the existence of optimal mechanisms from these lemmas. For a complete proof, we refer to Online Appendix A.1. We refer to Online Appendix A.3 for proofs of Lemmas 2 and 3 and to Online Appendix A.3.3. for that of Proposition 1.

To conclude, we note restricting to point-mass distributions is reasonable in hindsight because point-mass distributions happen to be the right class of "worst-case" distributions in that for any dynamic mechanism, there exists a point-mass distribution for the agent against which the dynamic mechanism is forced to obtain a regret at least the minimax regret. See the following proposition; its proof is provided in Online Appendix A.3.3.

Proposition 2. *Suppose Assumptions 1 and 2 hold. For any incentive compatible dynamic mechanism $A \in \mathcal{A}$ and $\epsilon > 0$, there exists a point-mass distribution δ_θ such that* $\mathrm{Regret}(A, \delta_\theta, T) \geq \mathrm{Regret}(T) - \epsilon$.

4. Applications

In this section, we apply our results to a dynamic selling mechanism design problem where a seller sells independent goods sequentially over time and a principal-agent model with hidden costs in which a principal repeatedly contracts with an agent to produce at particular output levels. A third application to a repeated resource allocation problem without monetary transfers is presented, due to space considerations, in Section TR.5 of the technical report Balseiro et al. (2019).

4.1. Dynamic Selling Mechanism

Consider a repeated setting where the principal (i.e., seller) sells independent and identical items to a strategic agent (i.e., buyer) over T rounds and seeks to maximize the revenue; this is an extension of the seminal single-round model from Bergemann and Schlag (2008) to our robust, dynamic mechanism design framework. The items are being sold one by one sequentially and, in each round, the agent realizes his value (equivalently, his willingness to pay) for the current item. The agent's values are drawn from an underlying private distribution known only to him. The principal does not know the agent's private value distribution except that the agent's value is in the range $[0, 1]$.

In the language of the general model, the agent's shock is his private value for the item and the shock space is $\Theta = [0, 1]$. We assume that $\mathcal{F} = \Delta([0, 1])$, which implies that Assumption 1 holds. The outcome space is $\Omega = \{0, 1\} \times \mathbb{R}$ and an outcome is $\omega = (\hat{x}, \hat{p}) \in \Omega$ where $\hat{x}$ is the allocation and $\hat{p}$ is the payment, that is, whether the item is allocated to the agent and the payment the agent makes to the principal. Given an outcome $\omega = (\hat{x}, \hat{p})$, the agent's utility function is $v(\theta, \omega) = \theta \cdot \hat{x} - \hat{p}$ and the principal's utility function is $u(\theta, \omega) = \hat{p}$. Abusing notations, for single-round direct mechanisms, we use $x : \Theta \to [0, 1]$ and $p : \Theta \to \mathbb{R}$ to denote the interim rules mapping reported shocks to expected allocations and payments, respectively, and the pair (x, p) to represent a single-round direct mechanism when convenient.

Note that $\mathrm{OPT}(\delta_\theta, T) = T\theta$ for all $\theta \in \Theta$ because the principal can extract the full surplus of the agent and satisfy the IR constraint by charging the agent's value when his shock is constant. Because of the agent's participation constraint, the principal's revenue is at most the agent's surplus and, thus, $\mathrm{OPT}(F, T) \leq T\mathbb{E}_{\theta \sim F}[\theta] = \mathbb{E}_{\theta \sim F}[\mathrm{OPT}(\delta_\theta, T)]$ for every distribution $F \in \mathcal{F}$. Therefore, Assumption 2 holds and Theorem 1 applies. We next show the minimax regret is T/e and an optimal dynamic mechanism is T repetitions of a randomized posted pricing mechanism to be specified below.

Proposition 3. *For revenue maximization in the dynamic selling mechanism design problem with one good, the minimax regret is T/e and an optimal solution is T repetitions of the randomized posted pricing mechanism S^* with price distribution Φ^* given by*

$$\Phi^*(p) = \begin{cases} 0, & \text{if } p \in [0, 1/e) \\ 1 + \ln p, & \text{if } p \in [1/e, 1] \end{cases},$$

such that the interim allocation and payment rules are $x^*(\theta) = \Phi^*(\theta)$ and $p^*(\theta) = [\theta - 1/e]_+$ for $\theta \in [0, 1]$.

For comparison, Amin et al. (2013) showed a lower bound of $T/12$ for the restricted class of dynamic posted pricing mechanisms; they considered a slightly different benchmark, but their results still hold in our setting (see section TR.6.1 of the technical report Balseiro et al. 2019). Their results would imply that static posted pricing mechanisms are asymptotically optimal in the restricted class. We consider more general dynamic mechanisms and determine an exactly optimal dynamic mechanism in this larger class with the performance that matches the improved lower bound of T/e exactly. To prove Proposition 3, it suffices to show that the randomized posted pricing mechanism S^* is a solution to the single-round minimax regret problem (3) and the corresponding minimax regret is $1/e$ via a saddle-point result, by Theorem 1. Its proof is deferred to Online Appendix B. An analogous result restricted to randomized posted pricing strategies exists due to Bergemann and Schlag (2008). Our single-round saddle-point result is for the slightly more general class of single-round direct IC/IR mechanisms and is still obtained with the same optimality structure and minimax regret value.[4]

Our results extend to multiple-goods settings (selling n goods in each round to an agent with additive, multidimensional valuations) and welfare maximization. In the case of welfare maximization, the minimax regret is zero because, using the Vickrey-Clarke-Groves mechanism (which reduces to allocating items for free for one buyer), it is possible to allocate efficiently without any prior beliefs on the agent's private information. In the case of revenue maximization with multiple goods, a robust optimal solution involves selling each item independently using the above mechanism S^* for a total regret of $(n/e)T$. Due to space considerations, we defer the analysis of these models to Sections TR.4.1 and TR.4.2 of the technical report Balseiro et al. (2019).

4.2. Principal-Agent Model with Hidden Costs

We consider a repeated principal-agent problem that captures various applications such as retail franchising, labor contracts, and procurement contracts. Similar to revenue maximization in the dynamic selling problem, we show that the minimax regret is linear in T and an optimal mechanism is static and repeats a single-round mechanism. Due to nonlinearity in the problem, our analysis is more involved.

More formally, the principal repeatedly contracts with the agent to produce output on his behalf and obtains revenue $R(\hat{q})$ when the agent produces $\hat{q}$ units of output, which is publicly observable. A contract specifies a payment $\hat{p}$ from the principal to the agent as a function of the number of output units $\hat{q}$. The agent has a private marginal production cost $\theta \in [\underline{\theta}, \bar{\theta}]$ where $0 < \underline{\theta} < \bar{\theta} < \infty$ which is assumed to be independently and identically distributed according to distribution F across the rounds. The agent observes his private cost and then decides on the production level $\hat{q}$ in each round. When he produces $\hat{q}$ units of output and receives a payment $\hat{p}$, his utility is $\hat{p} - \theta \cdot \hat{q}$ where θ is his marginal cost for that round. The principal does not know the agent's private distribution F but only that realized costs are in the range $[\underline{\theta}, \bar{\theta}]$. In particular, $\mathcal{F} = \Delta([\underline{\theta}, \bar{\theta}])$ and Assumption 1 holds. We assume $R(x)$ is a strictly increasing, strictly concave function that is twice continuously differentiable on $(0, \infty)$ with $R(0) = 0$ and $\lim_{x \to 0} R'(x) = \infty$; for example, $R(x) = \sqrt{x}$.

In terms of our general model, the agent's shock is his marginal cost of production and $\Theta = [\underline{\theta}, \bar{\theta}]$. The outcome space is $\Omega = \mathbb{R}_+ \times \mathbb{R}$ and an outcome $\omega = (\hat{q}, \hat{p}) \in \Omega$ is a pair of the production level $\hat{q}$ and the payment $\hat{p}$. When the outcome is $\omega = (\hat{q}, \hat{p})$ in a round, the agent's utility function is $v(\theta, \omega) = \hat{p} - \theta \cdot \hat{q}$ and the principal utility function is $u(\theta, \omega) = R(\hat{q}) - \hat{p}$. Abusing notations, for single-round direct mechanisms, we use $q : \Theta \to \mathbb{R}_+$ and $p : \Theta \to \mathbb{R}$ to denote the interim rules mapping reported shocks to production levels and payments, respectively, and the pair (q,p) to represent a single-round direct mechanism when convenient.

We now discuss the first-best mechanism that the principal can implement when he knows the agent's shock in a round. Because monetary transfers are allowed, the principal would set payments so that the IR constraint of the agent binds. Denote by $\bar{q}(\theta) = \text{argmax}_{x \geq 0}\{R(x) - \theta \cdot x\}$ the optimal production level when the agent's shock is known. The first-best mechanism involves requesting the agent to produce $\bar{q}(\theta)$ units and paying the agent the minimum amount $\theta \cdot \bar{q}(\theta)$ that makes him indifferent between participating or not in the contract (see, e.g., Laffont and Martimort 2001). Let $\bar{R}(\theta) = \max_{x \geq 0}\{R(x) - \theta \cdot x\}$ be the first-best utility of the principal when the shock is known to be θ. Because of the assumptions on $R(\cdot)$, $\bar{q}(\theta)$ is uniquely defined such that $R'(\bar{q}(\theta)) = \theta$ and $\bar{R}(\theta)$ is a strictly decreasing convex function.

A similar reasoning to the dynamic selling problem yields that $\text{OPT}(\delta_\theta, T) = T\bar{R}(\theta)$ because, when the agent's shocks are constant and equal to θ, the principal can implement the above first-best solution. The known-distribution benchmark for distribution F can be bounded as follows:

$$\begin{aligned}
\text{OPT}(F,T) &\leq \sup_{A \in \mathcal{A}} \mathbb{E}_{\pi,\sigma}\left[\sum_{t=1}^{T} R(\hat{q}_t) - \theta_t \hat{q}_t\right] \\
&\leq \sum_{t=1}^{T} \mathbb{E}_{\theta_t}\left[\bar{R}(\theta_t)\right] = \mathbb{E}_{\theta \sim F}\left[\text{OPT}(\delta_\theta, T)\right],
\end{aligned}$$

where the first inequality follows from the agent's participation constraint, which implies $\mathbb{E}_{\pi,\sigma}[\sum_{t=1}^{T} \hat{p}_t - \theta_t \hat{q}_t] \geq 0$, the second from relaxing the IC constraint and optimizing pointwise over the shocks, and the equality because shocks are identically distributed. Therefore, Assumption 2 holds and Theorem 1 applies. We formally state the main result of the subsection as follows:

Proposition 4. *For the principal-agent model with hidden costs, the minimax regret of the multiround problem is cT for some constant $c > 0$ and an optimal solution is T repetitions of offering the menu of deterministic contracts $\{(q^*(\theta), p^*(\theta))\}_{\theta \in \Theta}$, which is a single-round direct IC/IR mechanism. The allocation rule is continuous and satisfies the differential equation characterization*

$$(q^*)'(\theta) = -\frac{\bar{q}(\theta)}{R'(q^*(\theta)) - \theta}, \text{for } \theta \in (\underline{\theta}, \kappa),$$

with boundary conditions $q^*(\underline{\theta}) = \bar{q}(\underline{\theta})$ and $q^*(\theta) = 0$ for $\theta \in [\kappa, \bar{\theta}]$, where κ is the smallest cost for which q^* equals *to 0 and is assumed to be $\bar{\theta}$ if q^* is positive over $[\underline{\theta}, \bar{\theta}]$, and the payment rule is given by*

$$p^*(\theta) = \theta \cdot q^*(\theta) + \int_{\theta}^{\bar{\theta}} q^*(x)\mathrm{d}x.$$

For Proposition 4, we prove the stated single-round direct IC/IR mechanism is an optimal solution to the single-round minimax regret problem (3) via a saddle-point result. Because the revenue function R is concave, no randomization is needed and we can restrict our search of an optimal solution to those single-round direct IC/IR mechanisms that can be described in terms of a menu of deterministic contracts $(q(\theta), p(\theta))$ for $\theta \in [\underline{\theta}, \bar{\theta}]$ where the contract terms are all deterministic without randomization. We refer to Online Appendix C.1 for further details on the single-round problem including the IC/IR constraints and to Online Appendix C.2 for the proof of Proposition 4.

5. Interpretation and Necessity of Assumptions 1 and 2

In this section, we provide counterexamples showing that direct static mechanisms are not necessarily optimal in the absence of Assumptions 1 or 2. We also include a richer discussion of the extreme-point convexity assumption; we provide an economic interpretation, establish general sufficient conditions under which it holds, and provide an approximate form of Theorem 1 under an approximate extreme-point convexity condition.

5.1. Suboptimality of Static Mechanisms Without Assumption 1

Assumption 1 requires that the principal must guard against point-mass distributions because these are feasible choices for nature for the agent's distribution. To explore the necessity of this assumption, we show a game for which Assumption 1 does not hold and in which repeating a single-round direct mechanism leads to a suboptimal regret. The game will involve a class of shock distributions for which the knowledge of the mean immediately constrains the variance.

We consider a dynamic selling problem as in Section 4.1 with one good when the shock space is $\Theta = \mathbb{R}_+$ and the class of distributions $\mathcal{F}$ is a *scaled family* (see, e.g., Casella and Berger 2002, Definition 3.5.4). For a fixed cumulative distribution function G over $\mathbb{R}_+$, which is assumed to be known by the principal, we consider the parametric class of distributions $\mathcal{F}_G = \{F(\theta;\tau) = G(\theta/\tau)$ for some $\tau \in [0,1]\}$ where the scale parameter τ is known by the agent but not by the principal. Denoting by γ_t a random variable distributed according to G, we have that shocks are multiplicatively separable in the sense that $\theta_t \overset{(d)}{=} \tau\gamma_t$ (i.e., equivalent in the distributional sense). The effect of the scale parameter τ is to contract the distribution G while maintaining the same shape. For example, when G is an exponential distribution with mean 1, $\mathcal{F}_G$ is the class of all exponential distributions with mean $\tau \in [0,1]$. Alternatively, when G is a lognormal distribution with mean 0 and standard deviation σ, $\mathcal{F}_G$ is the class of all lognormal distributions with mean $\log(\tau) \in (-\infty, 0]$ and standard deviation σ.

Note we have an ex-ante participation constraint. Because monetary payments are allowed, Proposition 8 (in Subsection 5.2) implies that Assumption 2 holds. Assumption 1, however, does not hold because not all point-mass distributions are feasible distributions in $\mathcal{F}_G$.

We characterize an optimal mechanism using the relax and verify approach of Kakade et al. (2013):

Proposition 5. *Let G be an arbitrary distribution over $\mathbb{R}_+$ with mean $\mathbb{E}[\gamma] > 0$. For revenue maximization in the dynamic selling problem with one good and the agent's distribution restricted to $\mathcal{F}_G$, the minimax regret is $(\mathbb{E}[\gamma]/e)T$ and an optimal mechanism initially screens the agent by the parameter τ in Round 0, and then in each round, allocates each item with probability $x^*(\tau)$ and charges $p^*(\tau)$ with (x^*, p^*) as stated in Proposition 3.*

Recall that we can decompose shocks as $\theta_t = \tau\gamma_t$ with γ_t drawn i.i.d. from G. We prove the above result by considering a relaxed environment in which γ_t are public and observable by the principal—the agent's only private information is the parameter τ. Leveraging our ex-ante participation constraint, we show the relaxed problem can be reduced to a single-round problem that can solved to optimality as in the proof of Proposition 3. We conclude by showing that the resulting mechanism is feasible and, thus, optimal for the original problem. We remark that the mechanism presented in the above result is ex-ante individually rational but not ex-post individually rational, that is, the agent's total utility might be negative with a positive probability. We conjecture that the minimax regret of $(\mathbb{E}[\gamma]/e)T$ can also be asymptotically attained with an ex-post individually rational mechanism.

Let $\mathrm{Regret}^{\mathcal{S}}(T) := \inf_{S\in\mathcal{S}^{\times 1}} \sup_{F\in\mathcal{F}_G} \mathrm{Regret}(S^{\times T}, F, T)$ be the optimal regret for the class of direct static mechanisms. Because $\mathcal{S}^{\times T} \subset \mathcal{A}$, we have that $\mathrm{Regret}(T) \le \mathrm{Regret}^{\mathcal{S}}(T)$. For a specific game where direct static mechanisms without screening are strictly suboptimal or, equivalently, $\mathrm{Regret}(T) < \mathrm{Regret}^{\mathcal{S}}(T)$, we let G be the exponential distribution with mean 1. In this case, we have that $\mathrm{Regret}(T) = T/e$ from the above proposition. On the other hand, the following proposition shows that $\mathrm{Regret}^{\mathcal{S}}(T) = (1 - 1/e)T$.

Proposition 6. *Let G be the exponential distribution with mean 1. For revenue maximization in the dynamic selling problem with one good and the agent's distribution restricted to $\mathcal{F}_G$, the minimax regret for direct static mechanisms* *is* $(1-1/e)T$ *and repeatedly offering a posted price of 1 is an optimal direct static mechanism.*

Therefore, when Assumption 1 does not hold, there can be a separation between the minimax regrets achievable by incentive compatible dynamic mechanisms and by direct static mechanisms. The proofs of the above propositions are provided in Online Appendix D.

5.2 Sufficient Conditions for Assumption 2

This subsection provides sufficient conditions for the extreme-point convexity requirement in Assumption 2. Its economic interpretation is deferred until the next subsection. All proofs are deferred to Online Appendix E.

It is helpful to replace the known-distribution benchmark $\mathrm{OPT}(F,T)$ with a more tractable object. The first-best (equivalently, full information) benchmark $\bar{u}(F)$, defined below, is the optimal performance attainable when incentive compatibility constraints are dropped and the mechanism is subject only to an ex-ante individual rationality constraint. In other words, this is the optimal principal utility in a surrogate single-round problem in which the agent's distribution F is commonly known, shocks are publicly observable, and the agent commits to participate before observing the shock. Relatedly, we also consider the measure $\mathbb{E}_{\theta\sim F}[\bar{u}(\delta_\theta)]$ which differs from $\bar{u}(F)$ because it contains an interim IR constraint (on the per-shock basis) rather than an ex-ante IR constraint. It is the optimal principal utility in the single-round problem with full information as in $\bar{u}(F)$, but where the agent can choose not to participate (or report PASS) after the shock is realized. For any distribution F, we formally define $\bar{u}(F)$ and the (derived) measure $\mathbb{E}_{\theta\sim F}[\bar{u}(\delta_\theta)]$ as follows:

$$\bar{u}(F) := \sup_{S\in\mathcal{S}^{\times 1}} \int_\Theta\int_\Omega u(\theta,\omega)\mathrm{d}S_\theta(\omega)\mathrm{d}F(\theta) \quad \text{s.t.} \int_\Theta\int_\Omega v(\theta,\omega)\mathrm{d}S_\theta(\omega)\mathrm{d}F(\theta) \geq 0$$

$$\mathbb{E}_{\theta\sim F}[\bar{u}(\delta_\theta)] = \sup_{S\in\mathcal{S}^{\times 1}} \int_\Theta\int_\Omega u(\theta,\omega)\mathrm{d}S_\theta(\omega)\mathrm{d}F(\theta) \quad \text{s.t.} \int_\Omega v(\theta,\omega)\mathrm{d}S_\theta(\omega) \geq 0,\ \forall\ \theta\in\Theta.$$

Because the agent's distribution F and the shock θ in Round 1 are publicly known in the full information setting, it suffices that the principal designs a direct mechanism that assigns an outcome distribution for each possible shock value. When scaled by T, $T\bar{u}(F)$ and $T\mathbb{E}_{\theta\sim F}[\bar{u}(\delta_\theta)]$ are optimal performances achievable by the principal in the multiround problem with full information, respectively, under the ex-ante IR constraint and under the per-round interim IR constraint.

The next proposition relates these two quantities to Assumption 2:

Proposition 7. *For any* $F\in\mathcal{F}$, *if* $\bar{u}(F) = \mathbb{E}_{\theta\sim F}[\bar{u}(\delta_\theta)]$ *then* $\mathrm{OPT}(F,T) \leq \mathbb{E}_{\theta\sim F}[\mathrm{OPT}(\delta_\theta,T)]$.

Building on this broader sufficient condition, the next proposition gives more specific sufficient conditions that cover all applications in Section 4 and the technical report Balseiro et al. (2019). First, Assumption 2 holds for games with monetary payments that enter linearly into the utility functions of the principal and agent. This is because, given a mechanism satisfies the ex-ante IR constraint, the principal can use monetary transfers to satisfy the interim IR constraints without changing the expected overall utilities. The second part observes that Assumption 2 holds when agent utilities are always nonnegative, because both the ex-ante and interim IR constraints are satisfied automatically. When monetary transfers are allowed, and the principal wants to implement outcomes that lead to negative utility to the agent, our result may require negative payments to make the agent participate.

Proposition 8. *Assume the game is such that either (1) the outcome space factorizes as* $\Omega = \Omega^0\times\mathbb{R}$ *and the utility functions can be written* $u(\theta,(\omega^0,p)) = u^0(\theta,\omega^0)+\alpha p$ *and* $v(\theta,(\omega^0,p)) = v^0(\theta,\omega^0)-\beta p$ *for all outcomes* $(\omega^0,p)\in\Omega^0\times\mathbb{R}$ *for some functions* $u^0:\Theta\times\Omega^0\to\mathbb{R}$ *and* $v^0:\Theta\times\Omega^0\to\mathbb{R}$ *and scalars* $\alpha\geq 0$ *and* $\beta>0$, *or (2)* $v(\theta,\omega)\geq 0$ *for all* $\theta\in\Theta$ *and* $\omega\in\Omega$. *Then,* $\bar{u}(F) = \mathbb{E}_{\theta\sim F}[\bar{u}(\delta_\theta)]$ *for all* $F\in\mathcal{F}$ *and, therefore, Assumption 2 holds.*

5.3. Economic Intuition for Assumption 2 and Trade Across Shocks

Whether $\mathrm{OPT}(F,T)$ is extreme-point convex depends on two opposing effects of shock uncertainty. First, greater uncertainty in the shock distribution leads to greater information asymmetry between the principal and agent. To elicit the agent's private information, the principal needs to concede information rents, leading to lower principal utility. This suggests shock uncertainty can be undesirable from the principal's point of view. The second, more subtle effect, which we call *trade across shocks*, sometimes allows the principal to implement a broader set of outcomes when there is greater shock uncertainty. Subject to the agent's participation, the principal can implement outcomes that are beneficial to him but unfavorable for the agent under some realizations of the shock by, in return, offering other more favorable outcomes for the agent under other realizations of the shock. This kind of trading off is possible when there is shock uncertainty and, in this sense, shock uncertainty can be beneficial for the principal. Whether Assumption 2 holds depends on which of these effects dominates.

To see these effects directly, we can equivalently write the inequality in Assumption 2 as

$$T\bar{u}(F) - T\mathbb{E}_{\theta\sim F}[\bar{u}(\delta_\theta)] \le T\bar{u}(F) - \mathrm{OPT}(F,T),$$

by noting $\mathrm{OPT}(\delta_\theta, T) = T\bar{u}(\delta_\theta)$ (see Proposition E.1 in Online Appendix E).[5] The right-hand side is the effect of information asymmetry, which is commonly recognized to be the difference between the first-best (without IC) and "second-best" (with IC) performances in the Bayesian version of the multiround problem where F is known to the principal but not the realized shocks. The left-hand side is the effect of trade across shocks, which is the difference between optimal performances under different participation options, that is, the ex-ante IR versus per-round interim IR constraints, in the full information version of the problem where both F and shocks are known to the principal.

In many games where the principal knows F but not the shocks, the principal may link decisions across rounds subject to the agent's participation constraint so that $\mathrm{OPT}(F,T)/T \to \bar{u}(F)$ as $T\to\infty$ (see, e.g., Fudenberg et al. 1994, Jackson and Sonnenschein 2007). Therefore, the information asymmetry effect is negligible when the number of time periods is large and the trade across shocks effect ends up being the determining factor of whether the extreme-point convexity holds. In fact, Propositions 7 and 8 are statements about when the left-hand side is exactly 0.

To further explain trade across shocks, we focus on the full information setting, that is, F and θ_t are observed publicly. Consider the game in Table 1. There are two shocks or states of the world, reflecting the volume of rain in a given farming season. The principal cultivates a crop that requires heavy rain, whereas the agent cultivates a crop that grows only when the rain is light. There are two possible outcomes: they do not interact and each earns a utility of 0, or they share, which yields a utility of −1 for whoever must share his season-appropriate crop and a utility of 2 for the recipient (with the magnitude difference reflecting diminishing marginal returns).

For simplicity, assume each state θ^i of the world is equally likely under F. Then, a simple calculation shows $\bar{u}(F) = 3/4$ by sharing in θ^1 and sharing with probability $1/2$ in θ^2. On the other hand, $\bar{u}(\delta_{\theta^1}) = 0$ and $\bar{u}(\delta_{\theta^2}) = 0$, because the principal prefers not to share in θ^1 and the agent prefers not to in θ^2. In particular,

Table 1. A Game (Ω, Θ, u, v) with Outcome Space $\Omega = \{\emptyset, \omega^1\}$, Shock Space $\Theta = \{\theta^1, \theta^2\}$, and Utility Functions u and v of the Principal and Agent in Matrix Representation

$u(\cdot,\cdot)$, $v(\cdot,\cdot)$	$\emptyset$	ω^1 = SHARE
θ^1 = LIGHT RAIN	0,0	2, −1
θ^2 = HEAVY RAIN	0,0	−1, 2

Note. The no-interaction outcome is denoted by $\emptyset$.

$\bar{u}(F) > \mathbb{E}_{\theta\sim F}[\bar{u}(\delta_\theta)]$ because the principal can implement a broader set of outcomes when there is shock uncertainty. This gap reflects that sharing is individually rational for the agent ex-ante, but may not be after the state of the world is observed, and that the principal strictly does better under a sharing scheme of his choice that the agent agrees to before observing shocks.

To tie back to Theorem 1 where the principal does not know F, Assumption 1 implies the principal has to account for the possibility that F is a point-mass distribution and, when shocks are constant over time, direct static mechanisms that do away with trade across shocks and satisfy the more stringent interim individual rationality constraint are optimal. When Assumption 2 holds, by protecting against point-mass distributions, the principal can hedge against any other distribution as regret is the largest for point masses. In other words, because there is no benefit from trading across shocks, direct static mechanisms are optimal when the principal does not know the distribution F.

Now suppose that the extreme-point convexity does not hold and a nondegenerate distribution is worst-case optimal. Because the trade across shock effect dominates, direct static mechanisms might be suboptimal for two reasons. First, as discussed in the farming example, the interim individual rationality constraint limits the set of outcomes implementable by the principal. Second, even if we relax the participation constraint to ex-ante individual rationality, a dynamic mechanism might be necessary to attain low regret. This follows because, to implement outcomes that are individually rational for the agent in expectation over his shocks, the principal needs a dynamic scheme to infer the agent's distribution of shocks. In Section 5.4, we exhibit a game where direct static mechanisms incur linear regret and provide a dynamic mechanism that attains sublinear regret by inferring the agent's distribution of shocks.

Finally, imagine introducing money to the game above. From Proposition 8, we know $\bar{u}(F) = \mathbb{E}_{\theta\sim F}[\bar{u}(\delta_\theta)]$, because the principal could satisfy the more stringent interim IR constraint by paying the agent to share if necessary. Put more simply, the principal could buy crops from the agent in light rain seasons and sell his own crops in heavy rain seasons. Money allows for economic interactions that are individually rational to the agent in every state of the world. When monetary transfers are infeasible, the principal can try to mimic them by guaranteeing the agent beneficial outcomes in other states of the world. Trade across shocks is this phenomenon where one mimics money by transferring utilities across different states of the world to incentivize the agent's participation.

5.4. On the Regret of Static Mechanisms Without Assumption 2

The previous subsection provides intuition that dynamics may be beneficial when Assumption 2 fails due to the possibility of trade across shocks. But because our discussion was focused on a full information setting, it remains to show that dynamics can be beneficial in games with partial information. Due to space considerations, all proofs of the following results are deferred to Online Appendix F.

On the positive side, we show that direct static mechanisms are still near-optimal when Assumption 2 nearly holds, in the sense that the gap $\mathrm{OPT}(F,T)-\mathbb{E}_{\theta\sim F}[\mathrm{OPT}(\delta_\theta,T)]$ is small for all $F\in\mathcal{F}$.

Theorem 2. *Suppose Assumption 1 holds. For any $\epsilon\geq 0$, if a mechanism $S\in\mathcal{S}^{\times 1}$ satisfies (1), then*

$$\mathrm{Regret}(S^{\times T},T)\leq \mathrm{Regret}(T)+\epsilon + \sup_{F\in\mathcal{F}}\{\mathrm{OPT}(F,T)-\mathbb{E}_{\theta\sim F}[\mathrm{OPT}(\delta_\theta,T)]\}.$$

On the negative side, we next show a game for which Assumption 2 does not hold and direct static mechanisms are suboptimal. For this game, we prove that direct static mechanisms suffer linear minimax regret but that implementing a dynamic mechanism leads to a sublinear minimax regret. The performance gap shows that dynamic or adaptive schemes can effectively take advantage of the history of past outcomes and reports by the agent when Assumption 2 does not hold.

Consider the game in Table 2. It is not difficult to show that Assumption 2 fails. Let $F=(f_1,f_2)$ be the agent's private distribution over Θ where the shock is θ^i with probability f_i for $i=1,2$. We assume $\mathcal{F}=\{(f_1,f_2): f_1+f_2=1, f_i\geq 0\}$ so that Assumption 1 holds. The next result characterizes the optimal performance achievable.

Proposition 9. *For the game in Table 2, we have $\mathrm{OPT}(F,T)=T\cdot\bar{u}(F)$ and $\bar{u}(F)=\min\{f_1,f_2\}$.*

For point-mass distributions, we have $\bar{u}(\delta_{\theta^1})=\bar{u}(\delta_{\theta^2})=0$. For example, when $F=(1/2,1/2)$, we have $\bar{u}(F)=1/2$ but $\mathbb{E}_{\theta\sim F}[\bar{u}(\delta_\theta)]=f_1\cdot\bar{u}(\delta_{\theta^1})+f_2\cdot\bar{u}(\delta_{\theta^2})=0$; that is, $\mathrm{OPT}(F,T)>\mathbb{E}_{\theta\sim F}[\mathrm{OPT}(\delta_\theta,T)]$.

Like the game in Table 1, in this game, the principal would like to sometimes implement outcomes that are worse for the agent than not participating. Because the game does not allow for monetary transfers, the principal can only do this by making desirable expected outcomes in future rounds contingent on the agent's participation—implementing what we have called trade across shocks. The main difficulty in establishing the next proposition is in showing that the principal can implement a form of trade across shocks without knowing the shock distribution or observing the realized shocks.

Proposition 10. *For the game in Table 2, a separation exists:*

1. *For every T, the minimax regret of direct static mechanisms that repeat a single-round direct IC/IR mechanism is* $\inf_{S\in\mathcal{S}^{\times 1}}\sup_{F\in\mathcal{F}}\mathrm{Regret}(S^{\times T},F,T)=T/2$.
2. *For every T, there exists an incentive compatible dynamic mechanism A such that* $\sup_{F\in\mathcal{F}}\mathrm{Regret}(A,F,T)=O((\ln T)^{1/2}T^{2/3})$.

The first part of the proposition shows that any repetition of a single-round direct IC/IR mechanism necessarily incurs a linear regret. Because of the structure of the game, single-round direct IC/IR mechanisms are severely limited in generating utilities for the principal. For the second part, we design a dynamic mechanism A with two phases. In the first phase, the principal implements some default mechanism that places positive probabilities on outcomes ω^1 and ω^2 when the reported shocks are θ^1 and θ^2, respectively, and induces the agent to truthfully report his per-round shocks, as the present disutility from misreporting overwhelms any potential future gains. Then in the second phase, the principal estimates the agent's distribution F from the reports in the first phase and implements a better-tuned mechanism, which is a version of the optimal ex-ante mechanism that is perturbed to account for the statistical errors introduced by estimating F with limited samples. Using standard concentration inequalities, we can balance the loss of offering suboptimal mechanisms in the first and second phases. By choosing the number of rounds in the first phase to grow sublinearly relative to the time horizon, we can show that the dynamic mechanism incurs a sublinear regret.

Although we do not provide details due to the space consideration, (1) we can still show a separation between direct static mechanisms and dynamic mechanisms when the utility functions of the principal and agent are bounded, that is, bounded entries in Table 2; and (2) we can show the class of sequential screening mechanisms, which ask the agent to report his distribution and then implement a static mechanism based on the screened information, performs better than the class of (naive) direct static mechanisms considered above but still obtains the minimax regret that is linear in the time horizon (approximately, $0.217812\cdot T$).

It is worth mentioning that, from Proposition 8, Assumption 2 holds in the modified version of the

Table 2. A Game (Ω,Θ,u,v) with Outcome Space $\Omega=\{\emptyset,\omega^1,\omega^2\}$, Shock Space $\Theta=\{\theta^1,\theta^2\}$, and Utility Functions u and v of the Principal and Agent in Matrix Representation

$u(\cdot,\cdot)$, $v(\cdot,\cdot)$	$\emptyset$	ω^1	ω^2
θ^1	0, 0	1, −1	0, −∞
θ^2	0, 0	1, −∞	0, 1

Note. The no-interaction outcome is denoted by $\emptyset$.

above game that allows for monetary transfers. In that case, trade across shocks is not needed to incentivize participation and Theorem 1 shows direct static mechanisms are optimal.

6. Extensions and Discussion

Our results can be extended in several directions. Due to space considerations, we describe these briefly below and provide details in the technical report Balseiro et al. (2019).

• **Multiplicative guarantees.** We prove analogous results in terms of multiplicative performance guarantees instead of regrets. Similar to the approximation and competitive ratios in the theoretical computer science literature, we consider the multiplicative performance guarantee that is the ratio of the principal utility and the optimal performance achievable, as in $\mathrm{PrincipalUtility}(A, F, T)/\mathrm{OPT}(F, T)$, and the principal's goal is to maximize the worst-case ratio. Under Assumptions 1 and 2, we can show that the multiround multiplicative guarantee is equal to an appropriately defined single-round multiplicative guarantee and an optimal mechanism, if it exists, is static in that it repeats a single-round mechanism.

• **Dual Perspective and Saddle-Point Theorems.** The multiround and single-round minimax regret problems can be viewed as sequential-move zero-sum games in which the principal first chooses a mechanism and then nature selects a worst-case distribution to maximize the principal's regret. We show that, under certain conditions, these problems are respectively equivalent to ones in which nature chooses first and then the principal optimizes his performance given nature's choice. These results provide a framework for establishing the existence of optimal mechanisms and explicitly characterizing them and, also, a direct connection between our robust formulation and a more classical Bayesian formulation in the multiround problem.

• **Alternative benchmarks.** We show our results still hold for other alternative multiround benchmarks that are considered in the learning literature. Instead of the optimal performance achievable $\mathrm{OPT}(F, T)$, we consider $T \cdot \bar{u}(F)$, which is a stronger benchmark by Proposition E.1 in Online Appendix E, and a weaker benchmark which naturally corresponds to the performance achievable by repeating the best fixed single-round direct IC/IR mechanism (i.e., the best fixed "action" in hindsight). The latter has been considered by Amin et al. (2013) and subsequent works.

• **Arbitrary shock processes.** Our results also apply in the general shock process setting where the agent's shocks can be serially correlated according to a stochastic process that is known to the agent but not to the principal. This is a natural generalization of the repeated i.i.d. setting considered thus far where the per-round shocks are drawn independently and identically from an underlying distribution. As the set of shock processes is more general, the multiround minimax regret problem is more challenging for the principal. Not surprisingly, the constant shock processes where the agent's shock is fixed over the whole time horizon are the corresponding counterparts of point-mass distributions, which are worst cases in the repeated i.i.d. setting.

• **Principal pessimism.** We consider a stronger notion of regret in which the agent plays a utility-maximizing strategy that is least favorable for the principal. Under this alternative tie-breaking possibility, the worst-case uncertainty that the principal faces is in both the agent's distribution and his utility-maximizing strategy. We can show our general result (Theorem 1) and those in Section 4 still hold with respect to this more robust notion of minimax regret.

• **Connections to maximin utility objective.** We discuss some connections to the maximin utility objective for revenue maximization in the dynamic selling problem with a single good (Section 4.1). Despite differences in the settings and objectives, our results with respect to the minimax regret objective and those in Carrasco et al. (2019) with respect to the maximin utility objective have similar analyses and solution structures. We show this is because both papers rely on essentially the same single-round saddle-point problem involving direct IC/IR mechanisms and show equivalence-type connections using saddle-point results.

• **More stringent participation constraints.** Our model as stated assumes an ex-ante participation constraint in the sense that the agent determines whether to participate or not in Round 0 while knowing his distribution but not the realization of future shocks. This participation constraint is standard in the dynamic screening literature (see, e.g., Courty and Li 2000). Our results extend to other, more stringent participation constraints that have been considered in the literature such as the dynamic individual rationality constraint (Kakade et al. 2013, Pavan et al. 2014) and the per-period individual rationality constraint (Krishna et al. 2013, Ashlagi et al. 2016, Balseiro et al. 2018). This is because the minimax regret can be achieved by repeating a single-round direct IC/IR mechanism, which naturally satisfies the latter participation constraints.

• **Discounting.** We assume no discounting in Theorem 1, but when the principal and agent discount future payoffs using the same discount factor $\gamma \in (0, 1)$, the same results would still hold with minimal changes and the minimax regret would be linear in the effective time horizon $T_\gamma := 1 + \gamma + \ldots + \gamma^{T-1}$.

7. Conclusion

In this paper, we proved false-dynamics results for a finite horizon setting where the principal and agent repeatedly play a game. Our results hold whenever the set of possible distributions for the agent includes all point masses and when the optimal performance achievable is extreme-point convex. The latter condition is satisfied by all games with linear dependence on monetary transfers or in which the agent utility function is always nonnegative. In particular, this includes the dynamic selling problem, the principal-agent model with hidden costs, and resource allocation without monetary transfers, and we determined the minimax regret and characterized an optimal dynamic mechanism that simply repeats a single-round mechanism in these applications. When either assumption does not hold, it is possible that a dynamic mechanism can outperform static mechanisms and we showed a separation in terms of performance between dynamic and static mechanisms for specific games. Furthermore, we showed our techniques extend in several directions.

For future research, it would be interesting to better understand the extreme-point convexity assumption and find a more general class of games where false-dynamics-type results hold. On the other hand, it would be also interesting to further explore where false-dynamics-type results do not hold and identify the class of games in which there is a separation between dynamic and static mechanisms, that is, dynamics strictly helps.

Other possible research directions include restricting the space of distributions and considering multiple agents. Point-mass distributions happen to be the right class of worst-case distributions in our analysis, and we have shown that false-dynamics-type results do not necessarily hold when we rule out point-mass distributions. We considered one strategic agent who is forward-looking and responds to the principal's mechanism. When there are multiple forward-looking agents, equilibrium considerations become important as the outcome may differ depending on whether the agents know each other's distribution or not, which, in turn, may affect the principal's optimal dynamic mechanism. A final interesting research direction is to study the design of robust dynamic mechanisms when the principal has access to samples from the agent's distribution of shocks.

Acknowledgments

The authors thank the area editor Prof. Johari, an anonymous associate editor, and three anonymous reviewers for providing a number of valuable comments that greatly improved the paper. The authors also thank Dirk Bergemann, Ying-Ju Chen, and the participants at the INFORMS Revenue Management & Pricing Conference, ACM EC Workshop on Learning in Presence of Strategic Behavior, and Cornell ORIE Young Researchers Workshop for valuable feedback.

Endnotes

[1] Commitment can be sustained, for example, by writing a contract that is enforced by a court of law, by reputation, or when planning horizons are short. Lack of commitment can be shown to lead to worse performance for the principal.

[2] It is possible to apply the revelation principle and restrict attention to direct mechanisms in which the agent reports his distribution and the realized shocks. Because a direct mechanism asks the agent to report his private information and nothing more, the mechanism should only be defined for shocks that have a positive probability under the reported distribution. This creates some complexities, which we sidestep by proving a weaker version of the revelation principle: under some additional assumptions, there exists an optimal direct static mechanism.

[3] Note that, in general, (2) is not equivalent to $\text{Regret}(1) := \inf_{A\in\mathcal{A}} \text{Regret}(A, 1)$ because direct static mechanisms have the more stringent interim participation constraint and they do not allow the principal to screen the agent in Round 0 based on his distribution. Under our assumptions, however, these two problems are equivalent.

[4] To see the connection, note that for any single-round direct IC/IR mechanism S, there exists a randomized posted pricing strategy with nearly matching interim allocation and payment rules, that is, over $[0, 1]$ except a set of measure 0. For example, we can interpret a suitable extension and modification of the interim allocation rule of S as the cumulative distribution function from which posted prices are randomly drawn.

[5] Recall $T\bar{u}(F)$ and $T\mathbb{E}_{\theta\sim F}[\bar{u}(\delta_\theta)]$ are optimal performances achievable by the principal in the multiround problem with full information, respectively, under the ex-ante IR constraint and under the per-round interim IR constraint.

References

Amin K, Rostamizadeh A, Syed U (2013) Learning prices for repeated auctions with strategic buyers. Burges CJC, Bottou L, Welling M, Ghahramani Z, Weinberger KQ, eds. *Advances in Neural Information Processing Systems*, vol. 26 (Curran Associates, Inc., Red Hook, NY), 1169–1177.

Amin K, Rostamizadeh A, Syed U (2014) Repeated contextual auctions with strategic buyers. Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ, eds. *Advances in Neural Information Processing Systems*, vol. 27 (Curran Associates, Inc., Red Hook, NY), 622–630.

Ashlagi I, Daskalakis C, Haghpanah N (2016) Sequential mechanisms with ex-post participation guarantees. *Proc. 2016 ACM Conf. Econom. Computat., EC '16* (Association for Computing Machinery, Maastricht, The Netherlands), 213–214.

Bakos Y, Brynjolfsson E (1999) Bundling information goods: Pricing, profits, and efficiency. *Management Sci.* 45(12):1613–1630.

Balseiro S, Kim A, Russo D (2019) On the futility of dynamics in robust mechanism design–technical report. Technical report, Columbia University. http://www.columbia.edu/˜srb2155/papers/futility_technical_report.pdf.

Balseiro SR, Mirrokni VS, Leme RP (2018) Dynamic mechanisms with martingale utilities. *Management Sci.* 64(11):5062–5082.

Baron DP, Besanko D (1984) Regulation and information in a continuing relationship. Inform. Econom. Policy 1(3):267–302.

Bergemann D, Schlag KH (2008) Pricing without priors. *J. Eur. Econom. Assoc.* 6(2-3):560–569.

Bergemann D, Schlag K (2011) Robust monopoly pricing. J. Econom. Theory 146(6):2527–2543.

Bergemann D, Castro F, Weintraub G (2017) The scope of sequential screening with ex post participation constraints. *Proc. 2017 ACM Conf. Econom. Computat. EC '17* (ACM, New York), 163–164.

Besbes O, Zeevi A (2009) Dynamic pricing without knowing the demand function: risk bounds and near optimal algorithms. *Oper. Res.* 57:1407–1420.

Börgers T, Krahmer D, Strausz R (2015) *An Introduction to the Theory of Mechanism Design* (Oxford University Press, Oxford, UK).

Carrasco V, Luz VF, Monteiro P, Moreira H (2019) Robust mechanisms: The curvature case. *J. Econom. Theory* 68(1):203–222.

Carrasco V, Luz VF, Monteiro PK, Moreira H (2018b) Robust mechanisms: the curvature case. *Econom. Theory.*

Carrasco V, Luz VF, Kos N, Messner M, Monteiro P, Moreira H (2018a) Optimal selling mechanisms under moment conditions. *J. Econom. Theory* 177:245–279.

Carroll G (2017) Robustness and separation in multidimensional screening. *Econometrica* 85(2):453–488.

Casella G, Berger RL (2002) *Statistical Inference*, vol. 2 (Duxbury Pacific Grove, CA).

Courty P, Li H (2000) Sequential screening. *Rev. Econom. Stud.* 67(4):697–717.

Fudenberg D, Levine D, Maskin E (1994) The folk theorem with imperfect public information. *Econometrica* 62(5):997–1039.

Golrezaei N, Javanmard A, Mirrokni V (2020) Dynamic incentive-aware learning: Robust pricing in contextual auctions. Preprint, submitted February 25, https://arxiv.org/abs/2002.11137.

Green E (1987) Lending and the smoothing of uninsurable income. *Contractual Arrange Intertemp Trade* 1:3–25.

Jackson MO, Sonnenschein HF (2007) Overcoming incentive constraints by linking decisions1. *Econometrica* 75(1):241–257.

Kakade S, Lobel I, Nazerzadeh H (2013) Optimal dynamic mechanism design and the virtual pivot mechanism. *Oper. Res.* 61(3):837–854.

Kanoria Y, Nazerzadeh H (2020) Dynamic reserve prices for repeated auctions: Learning from bids. Preprint, submitted February 18, https://arxiv.org/abs/2002.07331.

Kleinberg R, Leighton T (2003) The value of knowing a demand curve: Bounds on regret for online posted-price auctions. *Proc. 44th Annual IEEE Sympos. Foundations Comput. Sci. FOCS '03* (IEEE Computer Society, Washington, DC), 594.

Kocyigit C, Rujeerapaiboon N, Kuhn D (2018) Robust multidimensional pricing: Separation without regret. Preprint, submitted August 9, https://dx.doi.org/10.2139/ssrn.3219680.

Kos N, Messner M (2015) Selling to the mean. Preprint, submitted July 18, https://dx.doi.org/10.2139/ssrn.2632014.

Krähmer D, Strausz R (2015) Optimal sales contracts with withdrawal rights. *Rev. Econom. Stud.* 82(2):762.

Krishna RV, Lopomo G, Taylor CR (2013) Stairway to heaven or highway to hell: Liquidity, sweat equity, and the uncertain path to ownership. *RAND J. Econom.* 44(1):104–127.

Laffont J-J, Martimort D (2001) *The Theory of Incentives: The Principal-Agent Model* (Princeton University Press, Princeton, NJ).

Laffont J-J, Tirole J (1993) *A Theory of Incentives in Procurement and Regulation*, vol. 1, 1st ed. (The MIT Press, Cambridge, MA).

Mohri M, Munoz A (2014) Optimal regret minimization in posted-price auctions with strategic buyers. Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ, eds. *Advances in Neural Information Processing Systems*, vol. 27 (Curran Associates, Inc., Red Hook, NY), 1871–1879.

Mohri M, Munoz A (2015) Revenue optimization against strategic buyers. Cortes C, Lawrence ND, Lee DD, Sugiyama M, Garnett R, eds. *Advances in Neural Information Processing Systems*, vol. 28 (Curran Associates, Inc., Red Hook, NY), 2530–2538.

Pavan A, Segal I, Toikka J (2014) Dynamic mechanism design: A Myersonian approach. *Econometrica* 82(2):601–653.

Pınar MÇ, Kızılkale C (2017) Robust screening under ambiguity. Math. Programming 163(1):273–299.

Spear SE, Srivastava S (1987) On repeated moral hazard with discounting. *Rev. Econom. Stud.* 54(4):599–617.

Thomas J, Worrall T (1990) Income fluctuation and asymmetric information: An example of a repeated principal-agent problem. *J. Econom. Theory* 51(2):367–390.

Wilson R (1987) Game-theoretic analysis of trading processes. Bewley T, ed. Adv. Econom. Theory *Fifth World Congress* (Cambridge University Press, Cambridge, UK), 33–70.

Santiago R. Balseiro is the Stanton Associate Professor of Business at the Graduate School of Business, Columbia University. His primary research interests are in the area of dynamic optimization, stochastic systems, and game theory with applications in revenue management and internet advertising.

Anthony Kim is an applied scientist at Amazon working on Sponsored Products. His research interests include revenue management, mechanism design, algorithmic game theory and online algorithms/learning. He received his PhD in Computer Science from Stanford University. Prior to Amazon, he was a postdoctoral researcher in the Decision, Risk, and Operations division at the Graduate School of Business, Columbia University.

Daniel Russo is an assistant professor in the Decision, Risk, and Operations Division of Columbia Business School. His research lies at the intersection of statistical machine learning and sequential decision-making, and he contributes to the fields of online optimization and reinforcement learning.