OPERATIONS RESEARCH
Vol. 61, No. 4, July–August 2013, pp. 837–854
ISSN 0030-364X (print) | ISSN 1526-5463 (online)
http://dx.doi.org/10.1287/opre.2013.1194

Optimal Dynamic Mechanism Design and the Virtual-Pivot Mechanism

Sham M. Kakade
Microsoft Research New England, Cambridge, Massachusetts 02142, skakade@microsoft.com

Ilan Lobel
Stern School of Business, New York University, New York, New York 10012, ilobel@stern.nyu.edu

Hamid Nazerzadeh
Marshall School of Business, University of Southern California, Los Angeles, California 90089, hamidnz@marshall.usc.edu

We consider the problem of designing optimal mechanisms for settings where agents have dynamic private information. We present the virtual-pivot mechanism, which is optimal in a large class of environments that satisfy a separability condition. The mechanism satisfies a rather strong equilibrium notion (it is periodic ex post incentive compatible and individually rational). We provide both necessary and sufficient conditions for immediate incentive compatibility for mechanisms that satisfy periodic ex post incentive compatibility in future periods. The result also yields a strikingly simple mechanism for selling a sequence of items to a single buyer. We also show that the allocation rule of the virtual-pivot mechanism has a very simple structure (a virtual index) in multiarmed bandit settings. Finally, we show through examples that the relaxation technique we use does not produce optimal dynamic mechanisms in general nonseparable environments.

Subject classifications: optimal mechanism design; dynamic mechanisms; dynamic private information; online advertising; sponsored search.
Area of review: Games, Information, and Networks.
History: Received February 2012; revisions received September 2012, March 2013; accepted March 2013.

1. Introduction

We study the problem of designing optimal mechanisms for environments with dynamic private information and propose a mechanism that is profit maximizing in a class of environments that we call separable. In a separable environment, the valuation function of an agent can be decomposed as the product (or the sum) of a function of the agent's first signal and another function of the agent's future signals.

A typical separable environment is one where the agent's value function depends on two or more kinds of private information, some of which are known in advance by the agent, while the others are learned or evolve over time. One example of such an environment is the one that occurs in online advertisement auctions, where a publisher sells the space on her website to advertisers. A typical advertiser will have two distinct kinds of relevant private information: she will know her profit margin on each sale and, because sales will generally be performed on the advertiser's own website, she will also have private information on conversion rates (the fraction of ads displayed that turn into sales). Because the advertiser can be expected to know a priori what her profit margin is, but should only learn over time what her conversion rate is, this example constitutes a separable environment.

Our theory also applies to the field of supply chain contracting. Consider the case of a manufacturer of a perishable product that supplies one or more retailers, who then sell the product onwards to the general public. This is also an example of a separable environment since the retailer will typically know her profit margin per good sold in advance, but her (potentially nonstationary) demand will have to be learned over time.

The optimal mechanism we propose, the virtual-pivot mechanism, is quite intuitive—it combines ideas based on the "virtual value" formulation of Myerson (1981) for static revenue-optimal mechanism design and the dynamic "pivot" mechanism proposed by Bergemann and Välimäki (2010) for maximizing social welfare. The mechanism essentially maximizes an affine transformation of the social welfare, which corresponds to a certain virtual surplus. Furthermore, the mechanism satisfies strong (periodic ex post) notions of incentive compatibility and individual rationality.

One notable special case of our results is the setting with only one buyer. Namely, consider a setting where the mechanism at each period has one item to sell to a single buyer. The mechanism has a fixed production cost γ for the item. Under separability assumptions, the optimal mechanism in this setting has a surprising simple form (with a simple indirect implementation that we present later)—the mechanism offers the agent a "menu" of contracts, of the form $(p, M(p))$, to the agent. If an agent chooses a contract, she will be charged an up-front payment of $M(p)$

and afterwards the mechanism posts a price of $p > \gamma$ at each time step—the agent has the option to pay more up-front for cheaper prices in the future. Note that even if the agent's valuation is increasing (or decreasing) over time and the seller is fully aware of this fact, the optimal mechanism involves offering the item at all periods at a constant price p.

In the general solution with multiple buyers, the virtual-pivot mechanism still retains this flavor. Roughly speaking, each agent, based on her initial type, is assigned a certain weight function in an affine transformation of the social welfare that is maximized by the mechanism; see §4.1. The more the agent pays up-front, the higher her importance will be in the social welfare function (leading to more allocations to her in the future).

Our setting considers a mechanism that allows agents to report their type every round. In particular, this implies that they are able to *re-report* all of their historical private information that has bearing on the current and future values. Allowing re-reporting of private signals is a crucial step in obtaining periodic ex post incentive guarantees. Once we obtain periodic ex post incentive compatibility for all future periods, we are able to provide necessary and sufficient conditions for incentive compatibility at the first period. We directly show that these conditions are satisfied for our optimal mechanism.

Finally, we provide examples of how the standard relaxation approach to dynamic mechanism design will not succeed without adding certain assumptions, such as separability.

1.1. Related Work

Two natural objectives in the dynamic mechanism design are maximizing the long-term social welfare of all buyers (*efficiency*) and maximizing the long-term revenue or profit of a seller (*optimality*). With regards to maximizing the long-term social welfare, there are elegant extensions of the efficient (VCG) mechanism to quite general dynamic settings, including the dynamic pivot mechanism of Bergemann and Välimäki (2010) and the dynamic team mechanism of Athey and Segal (2007) (see also Cavallo et al. 2007, Bapna and Weber 2008, Nazerzadeh et al. 2013).

The literature on the dynamic revenue-optimal mechanism has been primarily focused on settings where the agents arrive and depart dynamically over time, but their private information remains *fixed*; see Vulcano et al. (2002), Pai and Vohra (2013), Gallien (2006), Said (2012), Gershkov and Moldovanu (2009), and Skrzypacz and Board (2010). Several of these papers, including the first and the last one, are motivated by a revenue management setting where the underlying problem is dynamic because of the arrival of customers over time, but the customers themselves don't learn new private information over time. In this setting, the mechanism designer faces a dynamic problem, but the incentive constraints of each of the agents are essentially static because agents do not obtain any "new" private information over the course of the mechanism. For surveys on dynamic mechanism design, see Bergemann and Said (2011), and Parkes (2007).

We consider a setting where the private information of the agents changes over time, a line of research that was pioneered by Baron and Besanko (1984) and Courty and Li (2000). The latter provide an optimal mechanism for an environment where agents have private information about the future distribution of their valuations. Akan et al. (2008) showed how the optimal sequential screening mechanism changes if buyers have information about the time they learn their valuations. Battaglini (2005) studies a setting with a single agent whose private information is given by a two-state Markov chain and shows that the optimal allocation converges over time to the efficient allocation. In contrast to the results in Battaglini (2005), in the setting we consider, the allocation distortion generated by the agents' initial private information does not disappear over time (for a more detailed discussion, see §4.1, also Zhang 2012, Boleslavsky and Said 2012). See Battaglini (2005, 2007) also for results on optimal dynamic mechanism design in the absence of dynamic commitment power.

A closely related work to ours is that of Ëso and Szentes (2007), who study a two-period model where each agent receives a signal at the first period and the seller can also allow each agent to receive an additional private signal at the second period. Under certain concavity and monotonicity conditions on the signals, they show that the optimal mechanism allows the agents to receive their second signals; however, agents do not obtain any rents from the fact that the second-period signal is private. They also propose a "handicap" auction for the case where the agents' valuations are given by the sum of the first- and second-period signals. We use similar ideas and show that for a broad class of environments, the seller is able to extract the information rent associated with all signals except the initial one, even if the seller does not control the agents' ability to obtain further private signals. However, as we show in §6, there exist dynamic settings where the seller cannot extract the entire information rent from future signals. We also note the work in Deb (2008), which provides an optimal mechanism in a setting with only one buyer where the value is Markovian in the previous value, among other technical conditions.

Another paper closely related to ours is by Pavan et al. (2011). Their work is concurrent and has been developed independently from ours. They provide an envelope theorem and associated necessary conditions for mechanisms to be optimal in fairly general dynamic settings. They also provide some sufficient conditions for optimality of dynamic mechanisms that neither encompasses nor is encompassed by ours. We compare our necessary and sufficient conditions for optimality with theirs in §4.4.

1.2. Organization

We organize our paper as follows. In §2, we formalize our model, define separability, incentive compatibility, and optimality of mechanisms. In §3, we discuss our approach for designing optimal mechanisms. In §4, we propose our mechanism and state our main optimality result. Special cases (including the setting with only one buyer) are considered in §5. Section 6 provides simple examples showing how the usual incentive constraints from static mechanism are insufficient for the dynamic case. It also shows that without our separability assumptions, the particular relaxation approach we take is insufficient. The online appendix contains all the proofs. Supplemental material to this paper is available at http://dx.doi.org/10.1287/opre.2013.1194.

2. The Model

In this section, we formalize our model and define concepts such as incentive compatibility and optimality of mechanisms.

2.1. The Dynamic Environment

We consider a discrete-time, δ-discounted infinite-horizon ($t = 0, 1, 2, \ldots$) model that consists of one seller and n agents (buyers). The seller decides upon an action a_t at each period t among the feasible set of actions $\mathcal{A}_t$, at a cost of $c_t(a^t)$ to the seller, where $a^t = (a_0, a_1, \ldots, a_t)$ represents all the actions taken by the mechanism up to time t.

At every period, each agent $i \in \{1, \ldots, n\}$ receives a private signal $s_{i,t} \in S_{i,t}$. In particular, we make the following assumption about the first signal $s_{i,0}$ throughout the paper:

Assumption 2.1. *For each agent i, $s_{i,0} \in [0, 1]$ is real valued and distributed according to F_i. Furthermore, assume that F_i is strictly increasing and has a density, which we denote by f_i.*

This first signal summarizes all the initial private information of the agent (which has bearing on her entire stream of valuations). Furthermore, for all $t \geqslant 1$, each agent also receives a private signal $s_{i,t} \in S_{i,t}$—here we are not concerned with whether or not these future signals are real (the set $S_{i,t}$ is arbitrary for $t \geqslant 1$).

The *type* of agent i at time t is the sequence of signals of the buyer i up to (and including) time t, which is denoted by $s_i^t = (s_{i,0}, \ldots, s_{i,t})$. The type provides a summary of all the agent's private information, which has bearing on all her current and future valuations. For notational convenience, we let vector $s^t = \{s_i^t\}_{i \in [n]}$ denote the (joint) types of all agents at time t. At each period t, agent i obtains value $v_{i,t}(a^t, s_i^t)$, which is a function of her type and the seller's past and current actions. We assume quasi-linear utilities and denote the payment of agent i at time t by $p_{i,t}$, so that the (instantaneous) utility of agent i at time t is given by $u_{i,t} = v_{i,t}(a^t, s_i^t) - p_{i,t}$. We also assume throughout the following regularity condition.

Assumption 2.2. *The partial derivative $\partial v_{i,t}(a^t, s_{i,0}, \ldots, s_{i,t})/\partial s_{i,0}$ exists for all i, t, a^t, and s_i^t, and it is bounded by $\bar{V} < \infty$.*

We now specify the stochastic process over the signals. The signal $s_{i,t}$ that agent i receives at time t may be correlated to her previous signals $s_{i,0}, \ldots, s_{i,t-1}$ and the past actions of the seller $a_0, \ldots, a_{t-1}$, but it is independent (conditionally on the seller's actions) of all signals of the other agents. Formally, the stochastic signal $s_{i,t}$ is determined by the stochastic kernel $K_{i,t}(s_{i,t} \mid a^{t-1}, s_i^{t-1})$. We make the assumption that the first signal is independent of the future signals:

Assumption 2.3. *For each agent i, the distribution of the initial signal $s_{i,0}$ is independent of the future signals $s_{i,t}$ for $t \geqslant 1$.*

Even under this assumption, importantly, the value of agent i at any future period ($t \geqslant 1$) may still be correlated with the signal $s_{i,0}$. Here, we only explicitly assume $s_{i,0}$ to be independent of the future—arbitrary dependencies among future signals are permitted.

We also assume that the mechanism has the ability to exclude agents from the system at time $t = 0$. That is, it can select a subset of the agents that will obtain no value (and will not make payments) at any period $t \geqslant 0$. The exclusion of an agent from the system does not impact the value obtained by the other agents if the mechanism still takes the same sequence of actions $a_1, \ldots, a_t$.

Assumption 2.4. *The set of feasible actions $\mathcal{A}_0$ at time $t = 0$ is equal to $2^{\{1,\ldots,n\}}$, that is, the set of all subsets of $\{1, \ldots, n\}$. If $i \notin a_0$, then agent i is excluded from the system, i.e., $p_{i,t} = 0$ and $v_{i,t}(a^t, s_i^t) = 0$ for all t, a^t, and s_i^t. No agent obtains immediate value from the choice of a_0, i.e., $v_{i,0}(a_0, s_{i,0}) = 0$ irrespective of whether $i \in a_0$ or not. Also, the value obtained by each agent does not depend on the exclusion of other agents. In addition, the cost incurred by the mechanism only depends on the actions, not on the excluded agents.*

The assumption implies that for any pair of actions a_0, a_0' in $\mathcal{A}_0$ such that $i \in a_0$ and $i \in a_0'$, the value $v_{i,t}(a_0, a_1, \ldots, a_t, s_i^t) = v_{i,t}(a_0', a_1, \ldots, a_t, s_i^t)$ for all t, $a_1, \ldots, a_t$, and s_i^t. Also, $c_t(a_0, a_1, \ldots, a_t) = c_t(a_0', a_1, \ldots, a_t)$ for all t—of course, exclusion of an agent may change the choice of the actions taken by the mechanism. The assumption that the agents do not obtain value at $t = 0$ is made without loss of generality and for simplicity of presentation. Nevertheless, the mechanism may charge the agents $p_{i,t} \neq 0$ at that time. The above assumption simplifies satisfying the participation constraints. For example, if an agent only obtains negative values from the actions, she would be excluded from the mechanism. Observe that if the actions taken by the mechanism correspond to allocations of items to agent, this assumption can be simply satisfied.

Throughout the paper, suppose Assumptions 2.1–2.4 hold.

2.2. Separability

We now define a class of environments for which we construct optimal dynamic mechanisms. To be able to construct such mechanisms, we need to assume some structure on how the agents' values relate to their signals. The next property specifies two natural relationships between the signals and the values.

PROPERTY 2.1 (FUNCTIONAL SEPARATION). *An environment satisfies functional separation if the value function of each agent is either multiplicatively or additively separable*:

- *The value function of agent i is* multiplicatively separable *if there exists functions uniformly bounded A_i and $B_{i,t}$ such that*:

$$v_{i,t}(a^t, s_i^t) = A_i(s_{i,0}) B_{i,t}(a^t, s_{i,1}, \ldots, s_{i,t}). \tag{1}$$

- *The value function of agent i is* additively separable *if there exists uniformly bounded A_i, $B_{i,t}$, $C_{i,t}$ such that*:

$$v_{i,t}(a^t, s_i^t) = A_i(s_{i,0}) C_{i,t}(a^t) + B_{i,t}(a^t, s_{i,1}, \ldots, s_{i,t}). \tag{2}$$

DEFINITION 2.1. We call an environment *separable* if Assumption 2.3 and Property 2.1 hold.[1]

Separability specifies specific structural forms in how an agent's initial signal relates to her value function. Specifically, it ensures that it relates to the value function at each period via either a multiplicative or an additive form.

A curious reader might wonder why we would specify such structural assumptions for the initial signal, but impose so little structure on how future signals are correlated or how they relate to the value function. The answer is that the initial signals are the agents' private information when contracting first occurs. Therefore, the seller will have to pay an information rent for the agents' initial signals, but might hope not to pay information rent for signals the agents do not yet possess when contracting happens. This kind of decoupling of information rents between initial and future signals is not always possible in nonseparable environments, as we illustrate in §6, but the fact that it is indeed doable in separable environments is one of the messages of our paper.

2.3. Applications of Separable Environments

We now describe some examples of separable settings where the theory we develop is applicable.

Online Advertising. In Internet advertising (sponsored search), online publishers sell the space on their webpages via auctions to advertisers. Typically, an advertiser places an ad in order to: first, draw a user to visit the advertiser's website (via a click on the displayed ad), and then, subsequently, have the user perform a desired transaction such as purchasing a product or subscribing to a mailing list (cf. Mahdian and Tomak 2007, Nazerzadeh et al. 2013, Agarwal et al. 2009). The value that an advertiser obtains from the display of an ad depends both on the "conversion rate" (the probability that the user who sees the ad will choose to click on it and subsequently perform the desired transaction) as well as the profit that the firm obtains when the user performs the aforementioned transaction.

We assume that advertisers privately know the profit they obtain per transaction but are uncertain about the conversion rates. For instance, consider a firm (e.g., Amazon, Barnes and Noble) that sells books online and, in order to attract customers, advertises on search engines. When a user searches for a newly released book, the firm a priori knows the profit margin of selling that book, but only learns the conversion rate over time. In our model, the profit margin of each sale is represented by $s_{i,0}$. The action a_t represents which ads are shown to a given user and, potentially, in which slot each ad is shown. Every time the ad is shown to a user, the advertiser would obtain more information, represented by $s_{i,t}$s, and updates her belief about probability of a purchase. Therefore, $v_{i,t}(s_i^t) = s_{i,0} \times \Pr[\text{purchase} \mid a^t, s_{i,1}, \ldots, s_{i,t}]$. In the case where the publisher has either a single slot or a set of slots of identical quality, the typical approach used to update the probability of purchase is the following: the firm starts from a Beta-distributed prior, which is parameterized by the number of successful $x_{i,t}$ and failed $y_{i,t}$ conversions, and updates one of these two parameters every time the ad is displayed by incrementing either the number of successes or the number of failures depending on whether a transaction occurred. In this case, $B_{i,t}(a^t, s_{i,1}, \ldots, s_{i,t}) = \Pr[\text{purchase} \mid a^t, s_{i,1}, \ldots, s_{i,t}] = x_{i,t}/(x_{i,t} + y_{i,t})$. Note that even in the simple case of a single ad slot and a Beta-distributed prior, the simplest representation of advertiser i's knowledge about its conversion rate at time t, the pair $(x_{i,t}, y_{i,t})$, is a two-dimensional quantity.

What we call conversion rate is sometimes decomposed into two terms: a "click-through" rate that represents the probability that a user will click on an ad and a "conversion rate" that captures the probability that a desired transaction occurs given that the ad was clicked. Whereas conversion rates are typically learned privately by the advertiser, both the search engine and the advertiser are generally able to observe clicks on ads. To accurately capture the simultaneous private learning of conversion rates and public learning of click-through rates, we need to slightly expand the model to incorporate public signals as well as private ones. This can be done by incorporating new signals $\tilde{s}_{i,t}$ that are observed by both advertiser i and the search engine. In this case, the value of a click would be represented by $v_{i,t}(s_i^t, \tilde{s}_i^t) = s_{i,0} \times \Pr[\text{click} \mid a^t, \tilde{s}_{i,1}, \ldots, \tilde{s}_{i,t}] \times \Pr[\text{purchase given click} \mid a^t, \tilde{s}_{i,1}, \ldots, \tilde{s}_{i,t}, s_{i,1}, \ldots, s_{i,t}]$. Even though we describe our model and results without such public signals in order to simplify the notation throughout the paper, all of our results are valid for this slightly extended model as well.

Consider now a different online advertising setting where advertisers learn over time the monetary value of users referred to them by the search engine. Assume that each user's worth to advertiser i is equal to $\theta_i + \varepsilon_i$, where θ_i needs to be learned over time (its prior is a Gaussian distribution with mean μ_i and standard deviation σ_i) and ε_i is a zero-mean shock with standard deviation κ_i. Suppose that at the time of contracting, the search engine has already sent $N_{i,0}$ users to advertiser i, and the sum of the monetary worth of these $N_{i,0}$ users constitutes advertiser i's initial private information $s_{i,0}$. Let $s_{i,t}$ be equal to the monetary worth of user $t-1$ if that user was allocated to advertiser i and 0 if that user was not allocated to advertiser i. Then, this problem can be formulated as an additively separable environment. By Bayesian statistics, the expected value of user t to advertiser i is

$$v_{i,t} = \frac{\sqrt{\sigma_i^2 + \kappa_i^2}(s_{i,0} + \sum_{t'=1}^{t} s_{i,t'}) + \sigma_i \mu_i}{\sqrt{\sigma_i^2 + \kappa_i^2}(N_{i,0} + \sum_{t'=1}^{t} a_{i,t'-1}) + \sigma_i},$$

where $a_{i,t'-1}$ is an indicator of whether user $t'-1$ was allocated to advertiser i and, therefore, the advertiser received a new signal at time t' about the average value of the users. The environment is additively separable because, given the actions of the mechanism a^t, the function above is a linear combination of initial signal $s_{i,0}$ and the future signals $s_{i,1}, \ldots, s_{i,t}$.

Supply Chain Contracting: Consider a manufacturer of a perishable good who supplies one or more retailers over time. The retailers face a competitive market and sell the good at a market price of ρ, but each retailer i has its own private marginal operating cost, denoted by γ_i. The production cost of the manufacturer is given by $c(\cdot)$. The action a_t of the manufacturer can be decomposed into $(a_{1,t}, \ldots, a_{n,t})$ and represents how many units are shipped to the retailer in period t. Without loss of generality, we assume there is no lead time and the retailer receives the shipped units immediately. Retailer i faces demand $d_{i,t}$ at each period t, and the demand she encounters is private information. The revenue obtained by a retailer at period t is thus $v_{i,t}(\gamma_i, d_i^t) = (\rho - \gamma_i) \times \min\{d_{i,t}, a_{i,t}\}$. That is, the seller can sell the minimum between the demand she observes and the number of units she has in stock. Since the goods are perishable, there is no inventory carryover or inventory costs. The term $\rho - \gamma_i$ is initial private information of the retailer and, thus, is represented by $s_{i,0}$ is our model. We do not assume that the demand is stationary or has any particular structural form. In particular, we can let the signal $s_{i,t}$ at time t contain information about both current demand $d_{i,t}$ and future demand $d_{i,t'}$ for $t' > t$. As such, this model can allow for the retailers to be able to better forecast future demand than the manufacturer.

2.4. Mechanisms, Incentive Constraints, and Optimality

A mechanism $\mathcal{M}(q, p)$ is defined by a pair of an allocation rule $q(\cdot)$ and a payment rule $p(\cdot)$. We let $\mathcal{Q}$ denote the set of all allocation rules. By the Revelation Principle (cf. Myerson 1986), without loss of generality, we focus on (dynamic) direct mechanisms.[2] We assume the seller has full dynamic commitment power.

At each period t, each agent i makes a *report*, denoted by $\hat{s}_i^t$, of her type s_i^t. Using our standard shorthand notation, we denote the joint reports of all agents by $\hat{s}^t = \{\hat{s}_i^t\}_{i\in[n]}$. Note that because $s_i^t = (s_{i,0}, \ldots, s_{i,t})$ includes the set of all signals that each agent has received, each agent *re-reports* all of their previous signals at every period. The report of an agent can be conditioned on the history, which we now specify.

The public history at time t, denoted by h_t, is the sequence of reports and actions of the mechanism until period $t-1$; namely, $h_t = (\hat{s}_0, a_0, \hat{s}^1, a_1, \ldots, \hat{s}^{t-1}, a_{t-1})$. The private history of agent i at time t, denoted by $h_{i,t}$, includes the public history and her current type (sequence of signals she received up to, and including, time t), i.e., $h_{i,t} = (s_{i,0}, \hat{s}_0, a_0, s_{i,1}, \hat{s}^1, a_1, \ldots, s_{i,t-1}, \hat{s}^{t-1}, a_{t-1}, s_{i,t})$.

The allocation and payment rules are functions of the public history at time t, h_t, and the reports of all agents at time t, $\hat{s}^t$. The allocation rule determines the action taken by the mechanism, and the payment rule determines the payment of each agent.

The reporting strategy of agent i, denoted by R_i, is a mapping from her private history $h_{i,t}$ to a report of her current type $\hat{s}_i^t$. Mechanism $\mathcal{M}$ and the reporting strategy profile $R = \{R_i\}_{i\in[n]}$ determine a stochastic process which is described in Figure 1.

We now define the incentive constraints of the mechanism. Denote the expected (discounted) future value of agent i under the (joint) reporting strategy R in mechanism $\mathcal{M}$ by:

$$V_i^{\mathcal{M},R} = \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t v_{i,t}(a^t, s_i^t)\right]$$

and the expected (discounted) future utility (of i under R in $\mathcal{M}$) as:

$$U_i^{\mathcal{M},R} = \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \big(v_{i,t}(a^t, s_i^t) - p_{i,t}\big)\right],$$

where the expectation is with respect to the stochastic process induced by the reporting strategy and the mechanism.

Figure 1. A generic mechanism.

At each period $t \geqslant 0$, the following occurs:

1. Each agent i receives her private signal $s_{i,t} \sim K_{i,t}(\cdot \mid a^{t-1}, s_i^{t-1})$.
2. Each agent i provides a report, $\hat{s}_i^t$, of her current type, $s_i^t = (s_{i,0}, \ldots, s_{i,t})$, as determined by her private history $h_{i,t}$. In particular, $\hat{s}_i^t = R_i(h_{i,t})$.
3. As a function of the public history, h_t, and the current reports, $\hat{s}^t$, the mechanism determines the action $a_t \in \mathcal{A}_t$ and the payments $p_{i,t}$ for each agent i. In particular, $a_t = q(h_t, \hat{s}^t)$ and the joint prices are $\{p_{i,t}\}_{i\in[n]} = p(h_t, \hat{s}^t)$.

Similarly, for the expected value and utility of agent i, conditioned on a private history $h_{i,t}$ and type of the other agents s^t_{-i}, we have:

$$V_i^{\mathcal{M},R}(h_{i,t}, s^t_{-i}) = \mathbb{E}\left[\sum_{\tau=t}^{\infty} \delta^\tau v_{i,\tau}(a^\tau, s_i^\tau) \,\middle|\, h_{i,t}, s^t_{-i}\right]$$

$$U_{i,t}^{\mathcal{M},R}(h_{i,t}, s^t_{-i}) = \mathbb{E}\left[\sum_{\tau=t}^{\infty} \delta^\tau (v_{i,\tau}(a^\tau, s_i^\tau) - p_{i,\tau}) \,\middle|\, h_{i,t}, s^t_{-i}\right].$$

Note that this expectation is well defined (even on private histories which have probability 0 under R), since the reporting strategies are mappings from *all* possible private histories of agent i (and we have conditioned on the public history and current joint type).

Roughly speaking, the notion of incentive compatible is one in which no agent wants to deviate from the truthful strategy, as long as all other agents are truthful. This involves a somewhat delicate quantification with regards to the history. Our (weaker and stronger) notions of incentive compatibility are identical to those in Bergemann and Välimäki (2010).

Definition 2.2 (Incentive Compatibility). Let $\mathcal{T}$ denote the (joint) truthful reporting strategy.

• Dynamic mechanism $\mathcal{M}$ is (Bayesian) *incentive compatible* (IC) if, for each agent i, truthfulness is a best response to the truthful strategy of other agents—precisely, if for each i and R_i,

$$U_i^{\mathcal{M},\mathcal{T}} \geqslant U_i^{\mathcal{M},(R_i,\mathcal{T}_{-i})}.$$

• Dynamic mechanism $\mathcal{M}$ is *periodic ex post incentive compatible* if, for each agent i and at any time t, truthfulness is a best response to the truthful strategy of other agents—precisely, if for each i and time t, reporting strategy R_i, private history $h_{i,t}$, and current type of the other agents s^t_{-i}:

$$U_{i,t}^{\mathcal{M},\mathcal{T}}(h_{i,t}, s^t_{-i}) \geqslant U_{i,t}^{\mathcal{M},(R_i,\mathcal{T}_{-i})}(h_{i,t}, s^t_{-i}). \tag{3}$$

Note that the (weaker) Bayesian notion of IC implies that the truthful reporting strategy is a best response from a private history that is generated under $\mathcal{T}$ with probability 1. In contrast, the (stronger) periodic ex post notion demands that the truthful strategy is a best response on *any* private history, even those that have probability 0 under $\mathcal{T}$ (e.g., those histories where agents misreported in the past). See Bergemann and Välimäki (2010) for further discussion.

The notion of individual rationality is one, where at the equilibrium, the agents choose to participate (as it demands that the agents' utilities be nonnegative). Precisely,

Definition 2.3 (Individual Rationality). Let $\mathcal{T}$ denote the (joint) truthful reporting strategy.

• Mechanism $\mathcal{M}$ is (Bayesian) *individually rational* (IR) if, for each agent i, the expected future utility under the truthful strategy is nonnegative, i.e., $U_i^{\mathcal{M},\mathcal{T}} \geqslant 0$.

• Mechanism $\mathcal{M}$ is *periodic ex post individually rational* if the expected future utility is nonnegative for each agent i and time t, private history, $h_{i,t}$, and joint type of the other agents s^t_{-i}, i.e., $U_{i,t}^{\mathcal{M},\mathcal{T}}(h_{i,t}, s^t_{-i}) \geqslant 0$.

The *expected profit* of a mechanism $\mathcal{M}$ is the discounted sum of all payments of the agents minus the cost of the actions

$$\text{Profit}^{\mathcal{M}} = \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \left(-c_t(a^t) + \sum_{i=1}^{n} p_{i,t}\right)\right] \tag{4}$$

under the (joint) truthful reporting strategy $\mathcal{T}$. The objective of the seller is to maximize this expected profit, subject to both the incentive compatibility and individual rationality constraints. Precisely,

Definition 2.4 (Optimality). A Bayesian individually rational and Bayesian incentive-compatible mechanism is *optimal* if it maximizes the expected profit among all Bayesian individually rational and Bayesian incentive compatible mechanisms.

Note that the optimal mechanism is only required to satisfys the weaker Bayesian incentive constraints. This definition of optimality guarantees that the mechanism obtains an expected profit higher than (or at least equal to) any other mechanism that is incentive compatible and individually rational. Ideally, we might hope for an optimal mechanism that also satisfies the stronger (periodic ex post) incentive constraints, which ensures truthfulness is a best response even if agents have deviated in the past. As we show, the mechanism we propose, the virtual-pivot mechanism, enjoys these stronger guarantees.

3. A Relaxation Approach

We now provide a methodology for optimal dynamic mechanism design. The relaxation approach we take is the standard one also used in Ëso and Szentes (2007), Deb (2008), and Pavan et al. (2011). The difficulty is in "unrelaxing," i.e., showing that a candidate for the optimal policy satisfies the more stringent dynamic IC constraints.

Here, we are able to provide both necessary and sufficient conditions for dynamic IC. In particular, the use of the periodic ex post notion of incentive compatibility is critical in this characterization.

3.1. Relaxing

In this section, we consider a simpler, yet closely related, problem where we can utilize known static mechanism design techniques to design an optimal mechanism—these techniques are also used in Ëso and Szentes (2007), Deb (2008), and Pavan et al. (2011). The idea is to relax the optimization problem (of finding the optimal mechanism) by only imposing certain incentive constraints that arise in a simpler version of the problem. Roughly speaking, we attempt to solve a (simpler) less-constrained optimization problem. The critical issue is in showing that the solution to this less-constrained problem is also the optimal solution for the original problem.

Definition 3.1 (Relaxed Environment). Consider an environment where only the initial type $s_{i,0}$ is private to each agent i, whereas all her future signals are observed by the mechanism. We define this to be the *relaxed environment* and refer to our original environment as the *dynamic environment*.

Whereas the mechanism in the relaxed environment has full information with regard to the agents signals from $t \geqslant 1$, note that $s_{i,0}$ may affect all the future values of the agent. Observe that any direct mechanism in the dynamic environment induces a mechanism in the relaxed environment in a natural way: for $t \geqslant 1$, simply use the agents actual signals $s_{i,1}, \ldots, s_{i,t}$ as well as the reported initial signal $\hat{s}_{i,0}$ as the reported type $\{\hat{s}_i^t\}$ (as the input to the allocation and payment rules of the mechanism).

The following lemma is a rather straightforward observation.

Lemma 3.1. *Let $\mathscr{E}$ be a dynamic environment and $\mathscr{E}^{\text{relaxed}}$ be the corresponding relaxed environment. We have that*:

- *If $\mathscr{M}$ is an incentive compatible and individually rational mechanism in $\mathscr{E}$, then it is an incentive compatible and individually rational mechanism in $\mathscr{E}^{\text{relaxed}}$.*
- *Let $R^\star$ be the optimal revenue in $\mathscr{E}^{\text{relaxed}}$. Suppose a* (*Bayesian*) *incentive compatible and individually rational mechanism $\mathscr{M}$ in $\mathscr{E}$ has revenue $R^\star$; then, $\mathscr{M}$ is optimal for both $\mathscr{E}$ and $\mathscr{E}^{\text{relaxed}}$.*

This lemma suggest a natural optimal mechanism design approach: first, find an allocation rule $q^\star$ of an optimal mechanism in the relaxed environment $\mathscr{E}^{\text{relaxed}}$; then determine if there exists a pricing rule for $p^\star$ such that: (1) the mechanism $(q^\star, p^\star)$ is IC and IR in the dynamic environment $\mathscr{E}$; (2) the expected revenue it achieves is $R^\star$. If such a pricing is possible, then $(q^\star, p^\star)$ is optimal in $\mathscr{E}$. In our separable environments, we show that this approach is applicable. Furthermore, in §6, we discuss the limitations of this approach, where we provide certain nonseparable environments for which the optimal revenue in $\mathscr{E}$ is strictly less than the optimal revenue in $\mathscr{E}^{\text{relaxed}}$.

Envelope and Revenue Lemmas. Since in the relaxed environment the only piece of private information for each agent i is $s_{i,0}$, using the standard approach from static mechanism design (see Myerson 1981, Milgrom and Segal 2002), we provide the following lemma.

Lemma 3.2 (Envelope Condition). *Suppose that the mechanism $\mathscr{M}$ is IC in the relaxed environment. Then for all i, $s_{i,0}$, and $s'_{i,0}$,*

$$U_i(s_{i,0}, s_{-i,0}) - U_i(s'_{i,0}, s_{-i,0}) = \int_{s'_{i,0}}^{s_{i,0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \frac{\partial}{\partial s_{i,0}} v_{i,t}(a^t, s_{i,0}, s_{i,1}, \ldots, s_{i,t})\big|_{s_{i,0}=z} \,\bigg|\, s_{i,0}=z, s_{-i,0}\right] dz, \tag{5}$$

where $U_i(s_{i,0}, s_{-i,0})$ is the utility of agent i under the truthful strategy in $\mathscr{M}$, where the initial types are $s_{i,0}$ for i and $s_{-i,0}$ for the other agents.

Again, using standard techniques from static mechanism design, we can use the envelope condition above to establish the profit of any IC mechanism in the relaxed environment.

Lemma 3.3 (Expected Profit). *Suppose that the mechanism $\mathscr{M}$ is IC in the relaxed environment. Then, the expected profit obtained by the mechanism, Profit$^{\mathscr{M}}$, is equal to*:

$$\mathbb{E}\left[\sum_{t=1}^{\infty} \delta^t \left(\sum_{i=1}^{n}\left(v_{i,t}(a^t, s_i^t) - \frac{1 - F_i(s_{i,0})}{f_i(s_{i,0})} \cdot \frac{\partial v_{i,t}(a^t, s_{i,0}, s_{i,1}, \ldots, s_{i,t})}{\partial s_{i,0}}\right) - c_t(a^t)\right) - \sum_{i=1}^{n} U_i^{\mathscr{M}}(0, s_{-i,0})\right], \tag{6}$$

where the expectation is taken over $s_{i,0}$ and $s_{-i,0}$.

This lemma can be used to derive a candidate for the optimal allocation rule: if we pick an allocation rule that maximizes the equation above and pick a payment rule that makes it both IC and IR, then we will have an optimal mechanism.

3.2. The Relaxed Environment and the Virtual Welfare

In the relaxed environment, we can use the standard techniques of static mechanism design (Myerson 1981, Milgrom and Segal 2002) to establish an upper bound on the profit of the optimal mechanism. The next lemma establishes that in separable environments, the profit of any IC mechanism is an "affine transformation" of the social welfare of the agents. The affine factors are given by the functions α and β in the lemma. Note that they only depend on the initial signals (and the actions of the mechanism) and do not explicitly depend on the signals from $t \geqslant 1$. This observation underlies our construction of the optimal mechanism.

Lemma 3.4. *Consider the relaxed environment and an incentive-compatible mechanism $\mathscr{M}$. Suppose the environment is separable* (*as in Definition* 2.1)*. Then, under the stochastic process induced by $\mathscr{M}$ and the truthful reporting strategy, the expected discounted sum of payments by each agent i is equal to*

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t p_{i,t}\right] = \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t (\alpha_i(s_{i,0}) v_{i,t}(a^t, s_i^t) + \beta_{i,t}(a^t, s_{i,0}))\right] - \mathbb{E}\left[U_i^{\mathscr{M},\mathscr{T}}(s_{i,0}=0, s_{-i,0})\right], \tag{7}$$

where the functions α_i and $\beta_{i,t}$ are given by:

- *For multiplicatively separable values,*

$$\alpha_i(s_{i,0}) = 1 - \frac{1 - F_i(s_{i,0})}{f_i(s_{i,0})} \frac{A'_i(s_{i,0})}{A_i(s_{i,0})},$$

$$\beta_{i,t}(a^t, s_{i,0}) = 0.$$

- *For additively separable values,*

$\alpha_i(s_{i,0}) = 1,$

$$\beta_{i,t}(a^t, s_{i,0}) = -\frac{1-F_i(s_{i,0})}{f_i(s_{i,0})} A_i'(s_{i,0}) C_{i,t}(a^t).$$

The lemma above yields a bound on the profit of the optimal mechanism for the relaxed environment. Recall that Lemma 3.1 established that the profit for the dynamic environment is bounded by the profit from the relaxed one. Combining these two lemmas and the fact that an IR mechanism must satisfy $U_i^{\mathcal{M},\mathcal{T}}(s_{i,0}=0) \geqslant 0$, we obtain the following profit bound.

Corollary 3.1. *Under the assumptions in Lemma* 3.4, *for both the relaxed and the dynamic environments, the Profit*$^{\mathcal{M}}$ *of any incentive compatible and individually rational mechanism* $\mathcal{M}$ *is bounded as follows*:

$$\mathit{Profit}^{\mathcal{M}} \leqslant \max_{q\in\mathcal{Q}} \mathbb{E}\left[\sum_{t=1}^{\infty}\delta^t\left(\sum_{i=1}^{n}(\alpha_i(s_{i,0})v_{i,t}(a^t, s_i^t) + \beta_{i,t}(a^t, s_{i,0})) - c_t(a^t)\right)\right], \tag{8}$$

where $\mathcal{Q}$ *is the set of all allocation rules.*

The bound above determines an upper bound on the profit of any optimal dynamic mechanism. This bound is obtained by the allocation rule of the optimal mechanism for the relaxed environment. It does not, however, immediately yield an optimal dynamic mechanism since it does not determine the payments for the dynamic setting. In the next subsection, we discuss how to "unrelax," that is, how to obtain a mechanism for the dynamic setting from the allocation rule that maximizes the bound above.

3.3. Unrelaxing

From the relaxed environment, we can find a candidate for an optimal allocation rule. The main challenge here is how to find a payment rule and show that such a mechanism satisfies dynamic IC constraints. It turns out that it is natural to break this into two stages.

The first step is understanding how to ensure IC for $t \geqslant 1$. Here, there seems to be no general methodology in the literature (note that we are not assuming any structure on the stochastic process for the signals s_t, for $t \geqslant 1$). Our approach involves going one step further and trying to insure periodic ex post IC for periods $t \geqslant 1$. Recent work by Bergemann and Välimäki (2010) shows how to guarantee periodic ex post IC in the context of maximizing social welfare. Our results make use of this, but to do so, a critical conceptual step is to allow agents to re-report their entire type at every period. This way, we are able to obtain periodic ex post IC for $t \geqslant 1$.

For $t = 0$, where $s_{i,0}$ is real valued, we explicitly characterize the necessary and sufficient conditions for dynamic IC based on the fact that we have a periodic ex post IC mechanism for periods $t \geqslant 1$. This is a key technical step in our proof.

Re-Reporting and Periodic Ex Post IC. Recall that each agent i reports her entire type $s_i^t = (s_{i,0}, \ldots, s_{i,t})$ at each period t, not just her most recent private signal $s_{i,t}$. At the first glance, it may seem that this re-reporting of past private signals is redundant. It might even seem problematic, because it allows agents to give conflicting reports of their histories of signals received.[3] However, there are a few reasons why this approach is quite natural, both conceptually and technically.

Re-reporting significantly simplifies the task of obtaining periodic ex post IC guarantees. It gives an opportunity for agents that have reported untruthfully in the past to correct their past misreports and, in this way, return to truthful reporting course. In fact, it is unclear how to obtain such a guarantee for a mechanism that does not allow re-reporting in a setting with the same generality as ours (recall that we allow the signals for periods $t \geqslant 1$ to be drawn from arbitrary sets). Re-reporting enables us to construct a periodic ex post IC mechanism because it creates a way for the agents to inform the mechanism that previously submitted information is false and that the mechanism should instead consider a different, resubmitted history of events.

Obtaining periodic ex post IC guarantees is important for two reasons: first, it makes it far more likely that agents will indeed behave in an incentive-compatible way. With such guarantees, the agents' best response will always be to truthfully report their signals, no matter the history of events. If we could provide only Bayesian IC guarantees, the agents would only want to be truthful if they believed everyone had been truthful in every period up to that point in time. Given that we are designing mechanisms for complex dynamic settings, it is highly desirable to have the agents have proper incentives irrespective of the history of events. Second, periodic ex post incentive compatibility serves to break the problem of designing mechanisms for dynamic settings into simpler, smaller problems. That is, if we know that mechanism is periodic ex post IC from period $t+1$ onwards, we know that the agent will not have a profitable multiperiod deviation that involves a misreport in period t and a subsequent period $t' \geqslant t+1$. No matter what the agent does in period t, her incentive will be to be truthful from period $t+1$ onwards due to the periodic ex post IC guarantees. Therefore, proving the mechanism is periodic ex post IC from period $t+1$ onwards also means the only potential profitable deviations for the agent at period t are single-period deviations at that period. Checking that the agent does not have such single-period profitable deviations is a much easier task than showing that the agent does not have complex multiperiod profitable deviations.

Moreover, once we build a mechanism with re-reporting that is periodic ex post IC, we can also convert it to another mechanism without re-reporting that is still Bayesian IC. Let $\mathcal{M}$ be a periodic ex post IC mechanism with re-reporting and consider the mechanism $\overline{\mathcal{M}}$ with the same allocation and payment rule as $\mathcal{M}$, but where each signal is only reported once. That is, each agent i only reports signal $s_{i,t}$

at time t, and the period $t' \geqslant t$ report of signal s_t is replaced in $\bar{\mathcal{M}}$ by the unique report $\hat{s}_{i,t}$. Then, this new mechanism is Bayesian IC. The reason is as follows: re-reporting extends the set of strategies (deviations) of the agents. Being truthful is a strategy that is available in both mechanisms $\mathcal{M}$ and $\bar{\mathcal{M}}$. If being truthful is a Nash equilibrium of the game with a larger set of strategies, then it must also be a Nash equilibrium of this game with a restricted set of strategies. Therefore, Bayesian IC is maintained when we remove re-reporting. We note that even if our goal is to construct a Bayesian IC mechanism where agents report their types only once, considering the expanded mechanism where agents re-report their signals is still a useful technique in proving incentive compatibility. The technique we present here for proving Bayesian IC by considering a mechanism with re-reporting is novel and markedly different than the standard approach in literature, where the typical approach is to either restrict the types to be Markovian or to assume a structure on the possible signals so that every possible misreport could be corrected by a future second misreported signal. The sponsored search application, for example, is one where misreports cannot always be corrected by a second misreport, as we argued in the paragraph above. The types are also not Markovian unless you include the profit from a conversion and the number of successful and failed conversions in the type, in which case the agents would be reporting in every period all those pieces of information, creating a mechanism with effective re-reporting.

Necessary and Sufficient Conditions for IC. In the previous subsection, we argued that re-reporting simplifies the task of constructing a periodic ex post IC mechanism. We postpone the discussion of how we can use re-reporting to actually construct a periodic ex post IC mechanism until §4.

For now, assume that a mechanism $\mathcal{M}$ is periodic ex post IC for all periods $t \geqslant 1$. That is, for any period $t \geqslant 1$, any agent i, private history $h_{i,t}$, types of other agents s_{-i}^t, and reporting strategy R_i, Equation (3) is satisfied. We now provide necessary and sufficient conditions for such a mechanism to be IC (at period $t = 0$).

Consider a subset of an agent's reporting strategies that we denote by $x' \to x$. Define $x' \to x$ as the reporting strategy in which the agent reports x' as her first type $s_{i,0}$ (at $t = 0$), and subsequently (re-)reports it as x in all future periods ($t \geqslant 1$). Furthermore, under the strategy $x' \to x$, all other signals $s_{i,t}$ (for $t \geqslant 1$) are truthfully reported. In other words, at $t = 0$, she initially reports $\hat{S}_{i,0} = x'$, and, for $t \geqslant 1$, she reports $\hat{s}_i^t = (x, s_{i,1}, s_{i,2}, \ldots, s_{i,t})$. In $x' \to x$, we also allow x' and x to be functions of $s_{i,0}$. For example, the truthful strategy $\mathcal{T}_i$ can be represented as $s_{i,0} \to s_{i,0}$.

The expected utility of agent i under mechanism $\mathcal{M}$ and reporting strategy $x' \to x$ given her initial type $s_{i,0}$ is $U^{\mathcal{M},(x'\to x,\mathcal{T}_{-i})}(s_{i,0})$. For notational convenience, we drop the explicit dependence on the mechanism and the other agents' playing the truthful strategy and denote this by

$$U^{x'\to x}(s_{i,0}) = U^{\mathcal{M},(x'\to x,\mathcal{T}_{-i})}(s_{i,0}). \tag{9}$$

Similarly, we define the expected value of agent i under strategy $x' \to x$, assuming other agents are truthful by:

$$V^{x'\to x}(s_{i,0}) = V^{\mathcal{M},(x'\to x,\mathcal{T}_{-i})}(s_{i,0}). \tag{10}$$

We also use the notation

$$U^{x'\to x}(s_{i,0}, s_{-i,0}) \quad \text{and} \quad V^{x'\to x}(s_{i,0}, s_{-i,0}),$$

when we condition on the initial types of the other agents $s_{-i,0}$.

Suppose the mechanism $\mathcal{M}$ is one that is periodic ex post IC for periods $t \geqslant 1$. Under such a mechanism, if agent i deviates at period $t = 0$, while all other agents are truthful, agent i's best response strategy at all future periods $t \geqslant 1$ is to reveal her true type. Therefore, if her true first type is $s_{i,0}$, then to verify if truthfulness is a best response, we only need to verify that the truthful policy provides more utility then all misreporting strategies of the form $s'_{i,0} \to s_{i,0}$. Therefore, if mechanism $\mathcal{M}$ is periodic ex post IC for periods $t \geqslant 1$, then it is also IC at $t = 0$ if, and only if, for any true type x and time 0 report x',

$$U_i^{x\to x}(x) \geqslant U_i^{x'\to x}(x).$$

Subtracting $U_i^{x'\to x'}(x')$ from both sides, we get the following characterization: the mechanism $\mathcal{M}$ is IC if, and only if, for all x and x',

$$U_i^{x\to x}(x) - U_i^{x'\to x'}(x') \geqslant U_i^{x'\to x}(x) - U_i^{x'\to x'}(x'). \tag{11}$$

Furthermore $\mathcal{M}$ is periodic ex post IC if the above holds where we condition on the other types $s_{-i,0}$. That is, the mechanism is periodic ex post IC if for all x, x', and $s_{-i,0}$,

$$\begin{aligned} &U_i^{x\to x}(x, s_{-i,0}) - U_i^{x'\to x'}(x', s_{-i,0}) \\ &\quad \geqslant U_i^{x'\to x}(x, s_{-i,0}) - U_i^{x'\to x'}(x', s_{-i,0}). \end{aligned} \tag{12}$$

These observations are useful in that it we can use envelope conditions to precisely characterize incentive compatibility in terms of the expected values of the agents. First, we obtain that periodic ex post IC for $t \geqslant 1$ implies the following lemma.

LEMMA 3.5 (PERIODIC EX POST IC). *Suppose that mechanism $\mathcal{M}$ satisfies the periodic ex post IC conditions for all $t \geqslant 1$. Then, for all x and x' in $[0, 1]$, we have*

$$U_i^{x'\to x}(x) - U_i^{x'\to x'}(x') = \int_{x'}^{x} \frac{\partial V_i^{x'\to z}(s)}{\partial s}\bigg|_{s=z} dz. \tag{13}$$

It is straightforward to show that the partial derivative exists and, for any x, y, and z, is given by

$$\left.\frac{\partial V^{x\to y}(s)}{\partial s}\right|_{s=z} = \mathbb{E}\left[\sum_{t=0}^{\infty}\delta^t \frac{\partial}{\partial s_{i,0}} v_{i,t}(a^t, s_{i,0}, s_{i,1}, \ldots, s_{i,t})\Big|_{s_{i,0}=s} \;\Big|\; s_{i,0}=z\right], \tag{14}$$

where the expectation is under joint strategy $(x \to y, \mathcal{T}_{-i})$ in $\mathcal{M}$ (see Lemma A.1 in the appendix).

The following lemma uses the characterization above to obtain both necessary and sufficient conditions for incentive compatibility (at $t=0$).

Lemma 3.6. (Necessary and Sufficient Conditions for IC). *Suppose that the mechanism $\mathcal{M}$ satisfies the periodic ex post IC conditions for all $t \geqslant 1$. Then, $\mathcal{M}$ is IC for all $t \geqslant 0$ if, and only if, both conditions below are satisfied*:

- (*Envelope Condition*) *For all x and x'*,

$$U_i^{x\to x}(x) - U_i^{x'\to x'}(x') = \int_{x'}^{x} \left.\frac{\partial V_i^{z\to z}(s)}{\partial s}\right|_{s=z} dz. \tag{15}$$

- (*Interval Dominance*) *For all x and x'*,

$$\int_{x'}^{x} \left.\frac{\partial V_i^{z\to z}(s)}{\partial s}\right|_{s=z} dz \geqslant \int_{x'}^{x} \left.\frac{\partial V_i^{x'\to z}(s)}{\partial s}\right|_{s=z} dz. \tag{16}$$

Furthermore, $\mathcal{M}$ is ex post periodic IC if and only if the previous two conditions are satisfied when we condition on every possible other initial types $\hat{s}_{-i,0}$.

The result above is analogous to the characterization of incentive compatibility in standard single-parameter settings, where an envelope condition and monotonicity are used to characterize IC (see Myerson 1981). The envelope condition above is a standard one, but interval dominance replaces monotonicity in a dynamic setting. It compares the utility obtained by the truthful strategy (left-hand side) with other strategies of the form $x' \to s_{i,0}$ (right-hand side), because these are the only plausible candidate strategies when the mechanism is ex post IC for periods $t \geqslant 1$.

4. The Virtual-Pivot Mechanism

We now present the virtual-pivot mechanism, which is an optimal dynamic mechanism in separable environments.

The key insight from §3.2 is that the profit of a dynamic mechanism is bounded by an affine transformation of the social welfare of the agents, where the affine parameters are given by the functions α_i and $\beta_{i,t}$ in Lemma 3.4.

We define an affine weight function through a pair of vectors $(\hat{\alpha}, \hat{\beta})$, such that $\hat{\alpha} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_n) \in \mathbb{R}^n$ and $\hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_n) \in (\mathcal{A} \times \mathbb{R})^n$, where $\mathcal{A}$ includes all possible action vectors a^t for any t. In particular, $\hat{\beta}$ is allowed to depend action a^t, so that $\hat{\beta}(a^t) = (\hat{\beta}_1(a^t), \ldots, \hat{\beta}_n(a^t)) \in \mathbb{R}^n$. For any $(\hat{\alpha}, \hat{\beta})$, time t, and vectors of actions a^t and types s^t, the weighted social welfare with respect to $(\hat{\alpha}, \hat{\beta})$ is defined as

$$W^{(\hat{\alpha},\hat{\beta})}(a^{t-1}, s^t) \triangleq \max_{q\in\mathcal{Q}} \mathbb{E}\left[\sum_{\tau=t}^{\infty}\delta^\tau \left(\sum_{i=1}^{n}(\hat{\alpha}_i v_{i,\tau}(a^\tau, s_i^\tau) + \hat{\beta}_i(a^\tau)) - c_\tau(a^\tau)\right) \Big|\, s^t, a^{t-1}\right], \tag{17}$$

where the max is over all the possible allocation rules. Using a standard dynamic programming argument, the weighted social welfare satisfies the following (Bellman) equations:

$$W^{(\hat{\alpha},\hat{\beta})}(a^{t-1}, s^t) = \max_{a_t\in\mathcal{A}_t} \mathbb{E}\left[\sum_{i=1}^{n}(\hat{\alpha}_i v_{i,t}(a^t, s_i^t) - \hat{\beta}_i(a^t)) - c_t(a^t) + \delta W^{(\hat{\alpha},\hat{\beta})}(a^t, s^{t+1}) \,\Big|\, s^t, a^{t-1}\right], \tag{18}$$

where s_i^{t+1} is the next (random) type when conditioned on s^t and a^t.

Note, however, that the affine parameters $(\hat{\alpha}, \hat{\beta})$ we need to use to achieve the bound from Corollary 3.1 are not numbers (or, in the case of β, functions of the sequence of actions), but functions of the first signal $s_{i,0}$ of each agent i. An important challenge in implementing an IC mechanism is eliciting $s_{i,0}$ in an incentive-compatible way in order to obtain the desired $(\hat{\alpha}, \hat{\beta})$. An important design choice in the virtual-pivot mechanism is to use the first report of $s_{i,0}$ to determine the affine parameters $(\hat{\alpha}, \hat{\beta})$ and maintain those affine parameters fixed for all periods, irrespective of future re-reports of $s_{i,0}$. We note that, at any period, only the initial reports and the current period reports are used by the virtual-pivot mechanism, so past reports that are inconsistent with current reports are effectively ignored by the mechanism (except for the initial reports, which permanently impact the affine parameters).

The virtual-pivot mechanism is presented in Figure 2. The mechanism consists of two stages:

- (Subscription Phase) At time 0, each agent i, reports her initial type, $\hat{s}_{i,0}$. Then, the mechanism assigns affine parameters $(\hat{\alpha}_i = \alpha_i(\hat{s}_{i,0}), \hat{\beta}_i(\cdot) = \beta_{i,t}(\cdot, \hat{s}_{i,0}))$ to each agent i, where the functions α_i and β_i are given in Lemma 3.4. Then, the mechanism excludes the agents whose expected discounted payments would be negative (or zero). If $p_i^\star(\hat{s}_0) \leqslant 0$ (see definition in Equation (24)), then $i \notin a_0$. Otherwise, agent $i \in a_0$ and pays $p_{i,0}(\hat{s}_0)$ (see definition in Equation (25)).
- (Allocation Phase) For $t \geqslant 1$, the virtual-pivot mechanism is equivalent to an affine dynamic pivot mechanism. The affine parameters are fixed and the mechanism solicits reports from the agents in order to choose actions that maximize the affinely transformed social welfare $W^{(\hat{\alpha},\hat{\beta})}$.

To gain some intuition, let us consider the multiplicative-separable case. Roughly speaking, an agent with a higher initial signal $s_{i,0}$ would be assigned a larger $\hat{\alpha}_i$. A larger $\hat{\alpha}_i$ increases the weight of the agent in the affine transformation, and hence increases the value obtained by the agent.

We discuss the allocation and payment rules in more details in §4.2. Before that, we present our main results.

4.1. Optimality

We make the following assumptions.

Assumption 4.1 (Monotone Hazard Rate). *Assume that $f_i(s_{i,0})/(1-F_i(s_{i,0}))$ is strictly increasing.*

Assumption 4.2. *Assume that*

- (*Multiplicative Case*). *If the value function of agent i is multiplicatively separable, then $A_i(s_{i,0})$ is strictly increasing, twice differentiable, and concave in $s_{i,0}$.*
- (*Additive Case*). *If the value function of agent i is additively separable, then $A_i(s_{i,0})$ is strictly increasing, twice differentiable, and concave in $s_{i,0}$. Also, $C_{i,t}(a^t)$ is positive for all $a^t \in \mathcal{A}$,*

The function $A_i(s_{i,0}) = s_{i,0}$ is an example of a function that satisfies Assumption 4.2. These assumptions imply that α_i is strictly increasing for multiplicatively separable value functions and that $\beta_{i,t}$ is differentiable and strictly increasing for additively separable value functions (see Lemma A.2).

Theorem 4.1 (Optimality). *Suppose that the environment is separable and that Assumptions 4.1 and 4.2 hold. Then, the virtual-pivot mechanism is optimal in both the relaxed and the dynamic environments. In addition, the virtual-pivot mechanism is periodic ex post individually rational and periodic ex post incentive compatible.*

The proof of this theorem is presented in §4.3.

Figure 2. The virtual-pivot mechanism.

(Subscription Phase). At time $t=0$, for each agent i,
She reports $\hat{s}_{i,0}$.
Let $\hat{\alpha}_i \leftarrow \alpha_i(\hat{s}_{i,0})$, $\hat{\beta}_i(a^\tau) \leftarrow \beta_{i,\tau}(a^\tau, \hat{s}_{i,0})$ for all $\tau \geqslant 1$ and $a^\tau \in \mathcal{A}_\tau$.
If $p_i^\star(\hat{s}_0) \leqslant 0$ (see Equation (24)), then $i \notin a_0$ (agent i is excluded).
If $p_i^\star(\hat{s}_0) > 0$, then let $i \in a_0$ and charge her $p_{i,0}(\hat{s}_0)$, see Equation (25).

(Allocation Phase). At each time $t = 1, 2, \ldots$.
Each agent i reports $\hat{s}_i^t$.
Let $a_t^\star$ be an action that maximizes $W^{(\hat{\alpha},\hat{\beta})}(a^{\star t}, \hat{s}^t)$, see Equation (19).
Let $m_{i,t}$ be the flow marginal contribution of agent i, see Equation (21).
The payment of each agent i is equal to
$p_{i,t}(\hat{s}_t) \leftarrow v_{i,t}(a^{\star t}, \hat{s}_i^t) - m_{i,t}/\hat{\alpha}_i$.

The assumptions above allow us to satisfy the dynamic IC condition from Lemma 3.6. For optimality of the mechanism in the relaxed environment, a weaker set of assumptions could potentially be sufficient.

The virtual-pivot mechanism is optimal for both the relaxed and dynamic environments, and the profit obtained by the mechanism, as well as the utility obtained by the agents, are identical in both environments. Therefore, the agents obtain no "information rent" for periods $t \geqslant 1$. That is, the agents are not able to obtain any benefit from the fact that signals $s_{i,1}, \ldots, s_{i,t}$ are private. This no-information-rent property was noted in a two-period model by Ëso and Szentes (2007), where the mechanism is able to control whether or not agents obtain a second private signal. Theorem 4.1 implies that the no-information-rent property holds even in infinite-horizon problems where the sellers have partial control (or even no control) over what private signals agents obtain over time (signals evolve according to a stochastic kernel $K_{i,t}(s_{i,t} \mid a^{t-1}, s_i^{t-1})$), as long as the environment is separable. We show in §6 that this property does not extend to general nonseparable settings.

Because there is no information rent for periods $t \geqslant 1$, there is no allocation distortion associated with signals $s_{i,t}$ for $t \geqslant 1$. The initial signal $s_{i,0}$, however, causes distortion from the efficient allocation at *every* period as if the mechanism design problem was a static one. To see this easily, consider a setting where each agent i has a multiplicatively separable valuation and $A_i(s_{i,0}) = s_{i,0}$, i.e., the value function of agent i is $v_{i,t} = s_{i,0} \times B_{i,t}(a^t, s_{i,1}, \ldots, s_{i,t})$. The virtual-pivot mechanism allocates to maximize the "virtual valuations" of

$$\left(s_{i,0} - \frac{1-F_i(s_{i,0})}{f_i(s_{i,0})}\right) B_{i,t}(a^t, s_{i,1}, \ldots, s_{i,t}).$$

That is, the first signal $s_{i,0}$ is replaced at every period by the virtual value $s_{i,0} - (1-F_i(s_{i,0}))/f_i(s_{i,0})$ of static mechanism design (see Myerson 1981). Our results contrast to the ones of Battaglini (2005) and Zhang (2012), where the allocation distortion is transient (it disappears as t grows). This is due to the fact that in our model the impact of the signal $s_{i,0}$ is permanent (in the multiplicatively separable case, the signal $s_{i,0}$ multiplies $B_{i,t}(\cdot)$ for all t), whereas the impact of $s_{i,0}$ is transient in these other papers.

Applications to Online Advertising. In the generalized-second price (GSP) auctions (cf. Edelman et al. 2007) that are the prevalent mechanisms currently in use for sponsored search auctions, advertisers are ranked by their bids, multiplied by a quality score. The quality score is typically an estimate of the click-through rate of the advertiser. The price determined in the auction is also divided by this quality score. Hence, a larger click-through rate increases an advertiser's probability of the allocation and reduces its payments. Our results suggest the following form of contracts for sponsored search: the search engine would offer a menu of contracts to advertisers. Each contract would consist of an up-front payment and a multiplicative weight.

The weight purchased by the advertiser would work in a manner similar to the quality score (and, typically, in conjunction to it). An advertiser who purchased a given weight would see its bids multiplied by this weight during the auction and would see its payments divided by this weight. Advertisers with higher conversion rates would have an incentive to buy higher (and more expensive) multiplicative weights. Overall, advertisers who value an impression more (an impression means that an ad is shown to a customer) would pay more up-front, but pay less per auction and see its ad displayed more often.

4.2. The Allocation and Payment Rules

We first discuss the allocation rule of the mechanism. At each time t, the mechanism chooses allocation $a_t^\star$, which maximizes $W^{(\hat\alpha,\hat\beta)}(a^{\star t-1},\hat s^t)$, whereas $a^{\star t-1}=(a_0^\star,\ldots,a_{t-1}^\star)$ represents the past actions of the mechanism. From Equation (18), we have

$$a_t^\star \in \operatorname*{argmax}_{\{a_t\in\mathcal{A}_t\}}\bigg\{\sum_{i=1}^n\big(\hat\alpha_i v_{i,t}\big((a^{\star t-1},a_t),\hat s_i^t\big)+\hat\beta_i(a^{\star t-1},a_t)\big) -c_t(a^{\star t-1},a_t)+\delta\mathbb{E}\big[W^{(\hat\alpha,\hat\beta)}((a^{\star t-1},a_t),s_i^{t+1})\,|\,s^t=\hat s^t\big]\bigg\}. \tag{19}$$

Note that only reports from two time periods (0 and t) are used to determine $a_t^\star$. That is, $\hat s_0$ is used to determine the affine parameters and $\hat s_t$ is used to determine the agents' types at period t. At time t, the mechanism does not use the agents' reports between times 1 to time $t-1$ (for the allocation or payments).

We now show how the payments are determined. We start from the payments $p_{i,t}$ for $t\geqslant 1$ and then use those to construct $p_{i,0}$. To make the mechanism incentive compatible, $p_{i,t}$ is determined such that the (instantaneous) utility of agent i at time t is proportional to her flow marginal contribution to the affinely transformed social welfare, denoted by $m_{i,t}$.

$$\begin{aligned} m_{i,t} &= W^{(\hat\alpha,\hat\beta)}(a^{\star t-1},\hat s^t)-\delta\,\mathbb{E}[W^{(\hat\alpha,\hat\beta)}(a^{\star t},s^{t+1})\,|\,s^t=\hat s^t,a_t^\star] \\ &\quad - W_{-i}^{(\hat\alpha,\hat\beta)}(a^{\star t-1},\hat s^t) \\ &\quad + \delta\,\mathbb{E}[W_{-i}^{(\hat\alpha,\hat\beta)}(a^{\star t}_{-i},s^{t+1})\,|\,s^t=\hat s^t,a^{\star t},a^\star_{-i,t}], \end{aligned} \tag{20}$$

where $W_{-i}^{(a,b)}$ is the affinely transformed social welfare obtained in the absence of agent i

$$\begin{aligned} &W_{-i}^{(\hat\alpha,\hat\beta)}(a^{t-1},s^t) \\ &\triangleq \max_{q\in\mathcal{Q}}\mathbb{E}\bigg[\sum_{\tau=t}^\infty\delta^\tau\bigg(\sum_{j:\,j\neq i}\big(\hat\alpha_j v_{j,\tau}(a^\tau,s_j^\tau)+\hat\beta_j(a^\tau)\big) - c_\tau(a^\tau)\bigg)\,\bigg|\,s^t,a^{t-1}\bigg], \end{aligned}$$

and $a^\star_{-i,t}$ is the action that maximizes $W_{-i}^{(\hat\alpha,\hat\beta)}(a^{\star t-1},s^t)$ at time t.

Equivalently, we have

$$\begin{aligned} m_{i,t} &= \sum_{j=1}^n\big(\hat\alpha_j v_j(a^{\star t},\hat s^{j,t})+\hat\beta_j(a^{\star t})\big) \\ &\quad - c_t(a^{\star t}) - W_{-i}^{(\hat\alpha,\hat\beta)}(a^{\star t-1},\hat s^t) \\ &\quad + \delta\,\mathbb{E}\big[W_{-i}^{(\hat\alpha,\hat\beta)}(a^{\star t}_{-i},s^{t+1})\,|\,s^t=\hat s^t,a^{\star t},a^\star_{-i,t}\big]. \end{aligned} \tag{21}$$

The payment by agent i at time t is then given by

$$p_{i,t}(\hat s_t)=v_{i,t}(q^{\star t},\hat s_i^t)-\frac{m_{i,t}}{\hat\alpha_i}. \tag{22}$$

In Bergemann and Välimäki (2010), the idea of such a payment based on flow marginal contributions was introduced and shown to establish incentive compatibility for the welfare-maximizing allocation rule (see also Roberts 1979). Similarly, the payments that we use (which are scaled versions of the flow marginal contributions) establish incentive compatibility for the affinely transformed welfare-maximizing allocation rule.

We now construct the payment at time 0. Consider the allocation rule $q^\star$ that maximizes the weighted social welfare conditioned on the reports at time 0, i.e.,

$$q^\star\in\operatorname*{arg\,max}_{q\in\mathcal{Q}}\mathbb{E}\bigg[\sum_{t=1}^\infty\delta^t\bigg(\sum_{i=1}^n\big(\hat\alpha_i v_{i,t}(q^t,s_i^t)+\hat\beta_i(q^t)\big) - c_t(q^t)\bigg)\,\bigg|\,s_0=\hat s_0\bigg], \tag{23}$$

where $q_t=q(h_t,s^t)$ and $q^t=(q_0,\ldots,q_t)$. We drop the (explicit) dependence of q_t on h_t and s^t to simplify the presentation. Note that if the agents are truthful, then $q^\star$ and $a^\star$ correspond to the same allocation rule. Define $p_i^\star(\hat s_0)$ as follows:

$$p_i^\star(\hat s_0)=V_i(\hat s_0)-\int_0^{\hat s_{i,0}}\frac{\partial V_i^{z\to z}(s_{i,0},\hat s_{-i,0})}{\partial s_{i,0}}\bigg|_{s_{i,0}=z}dz, \tag{24}$$

where

$$\begin{aligned} &\frac{\partial V_i^{z\to z}(s_{i,0},\hat s_{-i,0})}{\partial s_{i,0}}\bigg|_{s_{i,0}=z} \\ &\quad=\mathbb{E}\bigg[\sum_{t=1}^\infty\delta^t\frac{\partial v_{i,t}(q^{\star t},s_{i,0},s_{i,1},\ldots s_{i,t})}{\partial s_{i,0}}\bigg|_{s_{i,0}=z}\,\bigg|\,s_{i,0}=z,\ s_{-i,0}=\hat s_{-i,0}\bigg]. \end{aligned}$$

The value $p_i^\star(\hat s_0)$ is the payment of agent i in the relaxed environment, given by the envelope condition. If $p_i^\star(\hat s_0)\leqslant 0$, then the mechanism excludes agent i (that is, $i\notin a_0^\star$).

The total expected discounted sum of payments in the relaxed and dynamic environments must match in order to achieve our optimality bound. Therefore, $p_i^\star(\hat s_0)$ must be equal to expected discounted sum of payments from agent i. Hence, the payment of agent i at time 0 equals

$$p_{i,0}(\hat s_0)=p_i^\star(\hat s_0)-\mathbb{E}\bigg[\sum_{t=1}^\infty\delta^t p_{i,t}(s_i^t)\,\bigg|\,s_0=\hat s_0\bigg]. \tag{25}$$

4.3. Unrelaxing: Proof of Theorem 4.1

In this subsection, we present the three steps of the proof of Theorem 4.1. The proofs of the following lemmas are given in the appendix.

The first step is to show that the mechanism, if incentive compatible, does indeed yield the profit from the upper bound in Corollary 3.1. The argument used to prove this lemma is a standard one from Myerson (1981). We also show that the virtual-pivot mechanism is periodic ex post individually rational.

Lemma 4.1. *If the virtual-pivot mechanism is incentive compatible, then it is optimal. Moreover, it is periodic ex post individually rational at* $t = 0$.

The lemma below guarantees that under the virtual-pivot mechanism, it is always a best response for agents to report their types truthfully regardless of the history, at any time $t \geqslant 1$ (assuming that other agents will be truthful in the future but not necessarily in the past). This lemma follows the technique of Bergemann and Välimäki (2010), except that it maximizes an affine transformation of the social welfare, instead of the social welfare itself.

Lemma 4.2. *The virtual-pivot mechanism is periodic ex post incentive compatible and periodicic ex post individually rational for all periods* $t \geqslant 1$.

The lemma above not only rules out deviations at periods $t \geqslant 1$, but it also rules out combined deviations at period $t = 0$ and future periods. That's because if an agent deviates at period 0, she still wants to truthfully report her type at a future period (the mechanism is periodic ex post IC).

Therefore, we need only concern ourselves with period $t = 0$ deviations from the truthful strategy. The proof of Theorem 4.1 is completed by the following lemma.

Lemma 4.3. *Suppose the assumption of Theorem* 4.1 *hold. Then the virtual-pivot mechanism satisfies the conditions provided by Lemma* 3.6 (*i.e., Equations* (15) *and* (16)). *These conditions are satisfied for all agents conditioned on any initial type* $s_{-i,0}$ *of the other agents and, therefore, the mechanism is periodic ex post incentive compatible.*

This is a key technical result in our paper. Proving this lemma involves addressing the key difference between the dynamic and the static setting, as we explicitly show the conditions of Lemma 3.6 hold. The separability assumption is central here.

4.4. On Our Methodology

Although other papers in the literature (see Ëso and Szentes 2007, Pavan et al. 2011) also provide optimal mechanisms using the relaxation approach, we emphasize that our construction and results do *not* immediately follow from them. The key challenge we address in our paper is showing that the allocation rule generated by the relaxation has an associated payment rule that makes the mechanism IC and IR in the dynamic setting. Our solution requires a combination of using the re-reporting technique, with constructing payments based on Bergemann and Välimäki (2010) to obtain periodic ex post IC for periods $t \geqslant 1$, as well as proving IC (at $t = 0$) by using our characterization of IC under the assumption of periodic ex post IC for $t \geqslant 1$.

Furthermore, we show in §6 that the relaxation approach does not work in every setting. In fact, the second example provides a simple dynamic environment in which the usual notions of monotonicity hold for the optimal allocation in the relaxed environment, and yet, this same allocation rule is not optimal in the dynamic environment (clearly showing how static notions of monotonicity are insufficient). Although we are not able to address the challenging problem of explicitly characterizing the necessary and sufficient properties of an environment for which this relaxation approach will succeed, we do provide environments in which both the relaxation approach fails and various assumptions of our separable environment are violated. Roughly speaking, these show that at least some variant of our assumptions are required for the relaxation approach to be successful.

Closely related to our work is the paper by Pavan et al. (2011), which concurrently and independently also develop a methodology for optimizing dynamic mechanisms. They construct an envelope theorem for dynamic environments and use it to provide necessary conditions for an optimal mechanism. Though quite general, their framework does not encompass ours (for example, they assume that all signals are real valued, whereas our work allows for signals in arbitrary sets for periods $t \geqslant 1$). The more delicate issue in designing optimal dynamic mechanisms is obtaining sufficient conditions for optimality, which requires creating a payment rule and proving it makes the mechanism incentive compatible. The mechanism we propose and the one proposed by Pavan et al. (2011) constitute two different mechanisms that are incentive compatible in different sets of environments. In particular, these two sets of environments do not encompass each other.

Our work focuses on a separability condition that allows for the design of optimal mechanisms. By separability, we mean that the first signal is independent from future signals and is related to the value function via a multiplicative or additive structure. Pavan et al. (2011) consider a notion of separability that is different from ours and that imposes several restrictions that do not apply to our work. Their definition of separability excludes our multiplicatively separable utility functions, which form the basis of our applications to online advertising and supply chain contracting. Their separable environments also require that the mechanisms' actions not affect the evolution of the agents' private information (see Assumption F-AUT). Having the mechanism's actions affect the evolution of types is important for our applications: in our sponsored search example, the advertiser should only learn about its conversion rate when its ad is displayed; similarly, in our supply chain contracting example, demand learning should not occur when the firm's inventory is too low. The implementation they propose is quite different from ours and relies on types being Markovian and real valued and the agents being able to correct past underreported signals by overreporting future ones. To ensure that agents can indeed correct a past misreport, they make stochastic dominance assumptions on the agents' types that we do not make (see Assumption F-FOSD). By establishing optimal mechanisms under two different sets of assumptions, our paper and Pavan et al. (2011) complement each other in the overall mission of finding settings where we can design optimal dynamic mechanisms.

5. Special Cases of the Virtual-Pivot Mechanism

In this section, we show that the virtual-pivot mechanism can be simply implemented in some natural special cases where it enjoys additional guarantees. First, we present an indirect implementation of the mechanism in an environment with a single agent. Then, we look at environments where the evolution of the types of the agents is either fully dependent or fully independent of the actions of the mechanism.

5.1. The Optimal-Contracting Mechanism for a Single Agent

We now consider the case where there is only a single agent. In this case, the optimal mechanism can be implemented as a remarkably simple *indirect* mechanism.

In particular, the indirect optimal-contracting mechanism is presented in Figure 3. The mechanism works as follows. The subscription phase is the only period at which the agent ever makes a report of her type. In particular, the agent just makes a report $\hat{s}_0$ of s_0.[4] In the posted-price-phase, the mechanism simply posts a price for every possible action; the agent decides upon the action; the agent pays the respective price for this action; the mechanism executes this chosen action. These prices may vary as a function of time because they depend on her previous purchases. After $t \geqslant 1$, the mechanism does not solicit reports from the agent.

Figure 3. The optimal-contracting mechanism for a single agent.

(Subscription Phase). At time $t = 0$,
The agent reports $\hat{s}_0$.
If $p^\star(\hat{s}_0) \leqslant 0$, then terminate the process (see Equation (24)).
Otherwise, charge the agent $p_0(\hat{s}_0)$ and continue (see Equation (25)).

(The Posted Price Phase). At each time $t = 1, 2, \ldots$.
The mechanism informs the agent of the price of each possible action, which is given by

$$p_t(a^t, \hat{s}_0) = \frac{c_t(a^t) - \beta_t(a^t, \hat{s}_0)}{\alpha(\hat{s}_0)}.$$

The agent chooses an action a_t, pays $p_t(a^t, \hat{s}_0)$, and the mechanism takes this action.

Corollary 5.1. *Suppose the assumptions of Theorem* 4.1 *hold and that there is only one agent. Then Optimal-Contracting is an optimal mechanism.*

In indirect mechanisms, we need to concern ourselves with what equilibrium we are implementing because agents are no longer simply reporting their types. The corollary above refers to the equilibrium where ties are broken as in the virtual-pivot mechanism.

To observe how simple the optimal-contracting mechanism is, consider a scenario where the mechanism is considering selling a stream of items to an agent. At each time period, the seller has two possible actions: allocate an item to the agent at a production cost $\gamma \geqslant 0$ or not (at no cost). The agents' valuation is multiplicative separable (hence, $\beta_t(a^t, \hat{s}_0) = 0$).

The optimal-contracting mechanism can be implemented as follows: the seller offers a family of contracts to the agent of the form $(p, M(p))$. The agent either leaves (and the process terminates) or she picks a price p. If the agent picks a price p she is immediately charged $M(p)$. At every period $t \geqslant 1$, the agent will offer to buy the item at the constant price p.

The value $M(p)$ the mechanism selects is

$$M(p) = p_0\left(\alpha^{-1}\left(\frac{\gamma}{p}\right)\right)$$

for each possible positive value of $p_0(s_0)$. In equilibrium, the agent will either leave (if $p^\star(s_0) \leqslant 0$) or will pick price $p = \gamma/\alpha(s_0)$.

This mechanism is optimal regardless of the value function of the agent, as long as it is multiplicatively separable. Even if the agent's value v_t is increasing or decreasing over time and the seller knows about it, it is still optimal for the seller to offer a family of contracts of the form $(p, M(p))$, which includes a constant price for every item ($t \geqslant 1$).

5.2. Controlled and Uncontrolled Environments

There are two natural extremes for how the stochastic process of the environment evolves. At one extreme is the fully *uncontrolled* environment, where the evolution of the agents' signals has no dependence on the action taken by the mechanism. Here, we show that the virtual-pivot mechanism enjoys a much stronger incentive-compatibility notion. At the other extreme is a *multiarmed bandit* process (which can be considered a fully controlled environment). Here, the type of an agent only evolves if the agent was allocated the item (and no evolution occurs otherwise) and the optimal allocation rule has a particularly simple form.

5.2.1. Fully Uncontrolled Environments. Define an *uncontrolled* environment to be one in which the stochastic process of each agent is independent of the actions taken by the mechanism, i.e., $K_{i,t}(s_{i,t} \mid a^t, s_i^{t-1}) = K_{i,t}(s_{i,t} \mid s_i^{t-1})$.

In this environment the allocation rule of the virtual-pivot mechanism is myopic, in that the mechanism's decision is to maximize the *instantaneous* weighted social welfare (as opposed to considering how this impacts future decisions). In particular, we have that:

$$\begin{aligned}\arg\max_{\{a_t \in \mathcal{A}_t\}} \mathbb{E}\left[\sum_{i=1}^{n} (\hat{\alpha}_i v_{i,t}(a^t, \hat{s}_i^t) + \hat{\beta}_i(a^t)) - c_t(a^t) \,\middle|\, \hat{s}^t, a^{t-1}\right] \\ = \arg\max_{\{a_t \in \mathcal{A}_t\}} \mathbb{E}\left[\sum_{i=1}^{n} \big(\hat{\alpha}_i v_{i,t}(a^t, \hat{s}_i^t) + \hat{\beta}_i(a^t)\big)\right. \\ \left. - c_t(a^t) + \delta W^{(\hat{\alpha},\hat{\beta})}(a^t, \hat{s}^{t+1}) \,\middle|\, \hat{s}^t, a^{t-1}\right].\end{aligned}$$

This is a straightforward corollary of the uncontrolled assumption.

Corollary 5.2 (A Dominant Strategy IC). *Suppose the assumptions of Theorem* 4.1 *hold. The virtual-pivot mechanism has the property that for every time step* $t \geqslant 1$, (*e.g., after time step* $t = 0$), *the truthful reporting strategy is a dominant strategy.*

This guarantee is immediate because each allocation from $t \geqslant 1$ is just instantly maximizing a social welfare function (and the action taken by the mechanism and the reports provided by the agents have no effect on the future evolution of signals). Hence, periodic ex post IC for periods $t \geqslant 1$ immediately implies ex post IC (and, hence, dominant strategies implementation) for periods $t \geqslant 1$. Hence, if agents knew their own (and other agents') past, present, and future signals, they would still report truthfully at all histories after $t = 0$. Note, however, that at period $t = 0$, the mechanism is still periodic ex post IC (not ex post IC).

5.2.2. Fully Controlled (Multiarmed Bandit) Environments. We now consider the setting where there is only one item to sell every round—so the action space for the mechanism at each period $t \geqslant 1$ consists of choosing which agent should receive the item (or choosing not to allocate the item). The environment now considered is one where the type of an agent evolves only if the mechanism takes an action. Namely, the type of an agent only changes when the mechanism allocates the item to the agent. We call this environment *controlled*; the underlying stochastic process corresponds to multiarmed bandits where each arm is mapped to an agent.

In a multiarmed bandit process, there is a "state" of each arm and this only evolves if the arm was "pulled." In our setting, *fully controlled* environment is one where if on any round $t - 1$ where agent i is not allocated the item, the signal $s_{i,t}$ is irrelevant. Precisely, we have that if i is not allocated at time $t-1$, then we have that: (1) all current and future values do not depend on $s_{i,t}$ and (2) the distribution of all future signals are independent of $s_{i,t}$. We also assume, for simplicity, that there are no costs associated with actions in the fully controlled setting.

A notable feature of this environment is that the optimal allocation is an index-based policy (a Gittins-type index, see Gittins 1989, Whittle 1982). Namely, we can assign a number to each agent, independent of the other agents, and the optimal allocation rule is to give the item to the agent with the highest positive index. In the fully controlled environment, the optimal allocation can be implemented using virtual indices.

Definition 5.1 (Virtual Index). For each agent i, the virtual index is defined as:

$$\mathcal{G}_i^{(\hat{\alpha},\hat{\beta})}(s_{i,t}) = \max_{\tau_i} \mathbb{E}\left[\frac{\sum_{t=t'}^{\tau_i} \delta^t (\hat{\alpha}_i v_{i,t'}(a^{t'}, s_i^{t'}) + \hat{\beta}_i(a^{t'}))}{\sum_{t=t'}^{\tau_i} \delta^t} \,\middle|\, s_{i,t}\right], \quad (26)$$

where the maximum is taken over all stopping times τ_i.

The optimal allocation rule is to give the item to the agent with the highest positive virtual index. The virtual index can be computed individually for each agent and, therefore, it decouples the n-agent problem into n single-agent problems.

The payments, however, cannot be computed separately for each agent because they depend on the externalities created by the agent receiving an item. The agents who do not receive an item at time t do not cause externalities and, therefore, do not make payments at time t (other than time $t = 0$). For the agent i that does get the item at time t, $W_{-i}^{(\hat{\alpha},\hat{\beta})}(a^{\star t}, \hat{s}^t) = \mathbb{E}[W_{-i}^{(\hat{\alpha},\hat{\beta})}(\hat{s}^{t+1}) \mid a^{\star t}]$. Hence, we obtain the following corollary.

Corollary 5.3 (The Virtual Index Mechanism). *Consider the fully controlled environment defined above and suppose the assumptions of Theorem* 4.1 *hold. The allocation rule of the virtual-pivot mechanism is to simply allocate to the agent with the highest virtual index. Moreover, for* $t \geqslant 1$,

$$p_{i,t}(\hat{s}^t) = \frac{1}{\hat{\alpha}_i}((1-\delta) W_{-i}^{(\hat{\alpha},\hat{\beta})}(a^{\star t}, \hat{s}^t) - \hat{\beta}_i(a^{\star t})).$$

To gain some intuition, consider the multiplicative-separable case. An agent with a higher initial type $s_{i,0}$ would be assigned a larger $\hat{\alpha}_i$. A larger $\hat{\alpha}_i$ increases agent i's virtual index and, therefore, increases the expected discounted value that agent i obtains. Moreover, she pays a lower payment at each period $t \geqslant 1$. However, for these privileges, she will be required to make a higher up-front payment (at $t = 0$).

6. Limitations of the Relaxation Approach

In this section, we provide examples where the optimal mechanisms in the dynamic and relaxed environments obtain different revenues. Our first example shows that if s_1 is correlated with the future signals, then the relaxation approach may fail. Our second example provides a simple, yet nonseparable, value function in which the relaxation approach fails. For more on how the relaxation approaches fails in general settings (that is, nonseparable environments), see the recent work by Battaglini and Lamba (2012).

The examples are two-period environments with one agent (e.g., future values can be considered to be 0, and we can set $\delta = 1$ without loss of generality). The agent receives signals s_0 and s_1 at times 0 and 1. At the end of the period $t = 1$, the mechanism takes an action $a \in \{0, 1\}$, corresponding to an allocation of an item. The agent obtains a value of $a \times v(s_0, s_1)$—no value is obtained at $t = 0$.

Correlated Signals. Suppose the value of the agent is equal to her second signal, namely, $v(s_0, s_1) = s_1$. Assume $s_0 \in [0, 1]$ and $s_1 \in [0, 1]$ are correlated. In the relaxed environment, the optimal mechanism is trivial: observe s_1, and take action $a = 1$, at the price equal to s_1. Hence, the optimal mechanism extracts the whole surplus, which is equal to $\mathbb{E}[s_1]$.

We now show that, under weak assumptions, the revenue of any dynamic mechanism that cannot observe the second signal is less than $\mathbb{E}[s_1]$. Consider an incentive-compatible and individually rational mechanism $\mathscr{M}$. Note that due to individual rationality constraints, a mechanism cannot extract more revenue than $\mathbb{E}[s_1 \times a^{\mathscr{M}}(s_0, s_1) \mid s_0] \leqslant \mathbb{E}[s_1 \mid s_0]$ from an agent of type s_0 where $a^{\mathscr{M}}(s_0, s_1)$ represents the mechanism's action (i.e., the probability of allocation). Thus, $\mathscr{M}$ can extract a revenue of $\mathbb{E}[s_1]$ only if $a = 1$ with probability 1.

On the other hand, if a mechanism chooses $a = 1$ with probability 1, then the expected payment at time $t = 0$, $\mathbb{E}[p_0 + p_1 \mid s_0]$, should be identical for all possible first-period types s_0 with probability 1, by Lemma 3.2 (if not, then the agent would misreport her type as the type with the minimum expected payment). Hence, the expected payment of the agent is less than or equal to $\inf_{s_0} \mathbb{E}[s_1 \mid s_0]$. Suppose s_0 and s_1 are correlated such that for a set κ of nonzero measure, if $s_0 \in \kappa$, then $\mathbb{E}[s_1 \mid s_0] < \mathbb{E}[s_1]$. In this case, the revenue of $\mathscr{M}$ is strictly less than $\mathbb{E}[s_1]$.

Nonseparable Value Functions. Now assume s_0 to be uniformly drawn from $[0, 1]$ and let s_1 be drawn independently and uniformly from the set $\{+, \times\}$. The value at time 1 is

$$v(s_0, +) = s_0 + c_+;$$

$$v(s_0, \times) = s_0 c_\times.$$

For all future times, assume the value is 0. Here, we assume c_+ is a constant greater than 1, and we later set $c_\times$ to be a large positive constant.

Note that this value function is of the form

$$v(s_0, s_1) = A(s_1)s_0 + B(s_1)$$

and does not satisfy our separability assumptions.

We observe that by Equation (6), there is a unique optimal allocation in the relaxed environment. This optimal allocation corresponds to the two static optimal auctions for the special cases where $s_1 = +$ and $s_1 = \times$. In particular, the allocation for $q(s_0, s_1 = +)$ is one that always allocates (because c_+ is greater than 1). The allocation for $q(s_0, s_1 = \times)$ occurs only if $s_0 \geqslant 0.5$. This allocation uniquely maximizes Equation (6) under the assumption that $U(0) = 0$.[5] To see this, note that for each setting of s_1, we have a static problem of optimal auction design with one item and one buyer. Furthermore, because the values are 0 at $s_1 = 0$, we have $U(0) = 0$.

It is interesting to note the following rather natural monotonicity properties:

- The value $v(s_0, s_1)$ is monotone (and linear) in s_0.
- The optimal (relaxed) utility is $U(s_0)$ is monotone in s_0.
- The future value $V(s_0)$ under the optimal allocation is monotone in s_0.

Nonetheless, we show that dynamic IC is more stringent and that the optimal revenue in the dynamic environment is less. Let $r^\star$ be this optimal revenue in the relaxed environment. Now observe that if $r^\star$ is achievable in the dynamic environment, then it must be due to this allocation rule—Equation (6) also specifies the expected payments in the dynamic environment. As a proof by contradiction, let us suppose that this allocation rule could be implemented in an IC manner in the dynamic environment.

Since the allocation does not change between 0 and 0.5, Lemma 3.2 implies:

$$U(s_0 = 0.5) - U(s_0 = 0) = \tfrac{1}{2}v(0.5, +) - \tfrac{1}{2}v(0, +).$$

Hence, the average revenue at $s_0 = 0$ is:

$$\begin{aligned} E\big[p_0 + p_1 \mid s_0 = 0\big] &= V(s_0 = 0) - U(s_0 = 0) \\ &= \tfrac{1}{2}v(0, +) - U(s_0 = 0) \\ &= \tfrac{1}{2}v(0.5, +) - U(s_0 = 0.5). \end{aligned}$$

Now consider the misreporting strategy R of using $\hat{s}_0 = 0$ when $s_0 = 0.5$ and then reporting $\hat{s}_1 = \times$ when $s_1 = +$ and reporting $\hat{s}_1 = +$ when $s_1 = \times$. Here, the agent obtains the item when $(s_0, s_1) = (0.5, \times)$ (since $(\hat{s}_0, \hat{s}_1) = (0, +)$ is reported, which leads to an allocation). The value under this strategy is

$$V^R(s_0 = 0.5) = \tfrac{1}{2}v(0.5, \times)$$

(since with a 1/2 probability the agent obtains $s_1 = \times$). Also, note that the distribution of misreports $\hat{s}_1$ is uniform under R, so that the expected payments under R at $s_0 = 0.5$ are identical to those at $s_0 = 0$. Hence,

$$\begin{aligned} U^R(s_0 = 0.5) &= V^R(s_0 = 0.5) - E[p_0 + p_1 \mid s_0 = 0] \\ &= \tfrac{1}{2}v(0.5, \times) - \tfrac{1}{2}v(0.5, +) + U(s_0 = 0.5) \\ &= \tfrac{1}{2}(0.5c_\times - 0.5 - c_+) + U(s_0 = 0.5). \end{aligned}$$

Thus, for sufficiently large $c_\times$, we have that this misreporting strategy obtains strictly greater utility than that of the truthful strategy. Furthermore, by a continuity argument, for a neighborhood $[0.5, 0.5 + \epsilon]$ this misreporting strategy will also provide strictly more revenue (since the allocation rule does not change above $s_1 \geqslant 0.5$). Thus, we have a contradiction—there is a misreporting strategy resulting in strictly greater (unconditional) expected utility.

7. Concluding Remarks

In this work, we propose an optimal dynamic mechanism, the virtual-pivot mechanism, for separable environments. Separability is a condition that is often satisfied when the agents have multiple different kinds of private information, some of which they know in advance and other that they learn over time. Separability arises in several different settings, from the world of online advertising to the problem of supply chain contracting.

Our methodology is as follows: we first find a candidate allocation rule by solving the mechanism design problem in a relaxed environment, as is standard in this literature. The key challenge we address is how to find a (dynamic) payment rule that makes this candidate allocation rule incentive compatible. Our solution methodology involves aiming for a bigger goal: finding a payment rule that makes the candidate allocation rule periodic ex post incentive compatible. We show that this is possible for periods after the initial one if we allow the agent to "re-report" their entire history of signals at each period. In particular, the payment rule we need is constructed by mapping the candidate allocation rule to an affine transformation of the social welfare function. We find necessary and sufficient conditions for incentive compatibility at the initial periods for mechanisms that satisfy periodic ex post incentive compatibility for periods after the first one. Finally, we show that the virtual-pivot mechanism satisfies these conditions and is, therefore, incentive compatible.

The virtual-pivot mechanism is quite simple and could be implemented in settings such as selling online advertisement (see §§2.2 and 4.1). The variant of this mechanism specialized to one-buyer settings, the optimal-contracting mechanism, is even simpler and shows that the structure of the optimal mechanism can be quite counterintuitive.

We show in §6 that this relaxation approach will not work in designing optimal mechanisms for general nonseparable settings. The precise extent to which our technique works in nonseparable settings and what methodology could be used in designing optimal mechanisms when the relaxation method fails are promising areas for future research.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/opre.2013.1194.

Endnotes

1. We do assume that Assumption 2.3 holds throughout the paper, but we state the definition above as a combination of Property 2.1 and Assumption 2.3 to clearly state that for an environment to be separable, the value function of each agent must satisfy both a functional and a statistical (independence of first signal) separation.
2. The Revelation Principle implies that an equilibrium outcome in any indirect mechanism can also be induced as an equilibrium outcome of an (incentive-compatible) direct mechanism.
3. See §4 for how the mechanism utilizes the (potentially incoherent) sequence of reports provided by the agents.
4. Observe that the subscription phase can be implemented in an indirect manner by offering a menu of contracts at time 0. However, for the simplicity of presentation, we assume the agent reports her initial type.
5. Again, technically, there is a family of maximizers that agrees with probability 1. The argument holds for any of these maximizers.

Acknowledgments

The authors thank Maher Said, Gregory Lewis, Daron Acemoglu, Susan Athey, Markus Mobius, Mallesh Pai, Andrzej Skrzypacz, Rakesh Vohra, the associate editor, and the referees for many insightful suggestions and helpful comments. All three authors thank Microsoft Research New England for its support.

References

Agarwal N, Athey S, Yang D (2009) Skewed bidding in pay per action auctions for online advertising. *Amer. Econom. Rev.: Papers Proc.* 99(2):441–447.
Akan M, Ata B, Dana J (2008) Revenue management by sequential screening. Working paper, Carnegie Mellon University, Pittsburgh.
Athey S, Segal I (2007) An efficient dynamic mechanism. Working paper, Stanford University, Stanford, CA.
Bapna A, Weber T (2008) Efficient dynamic allocation with uncertain valuations. Working paper, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.
Baron DP, Besanko D (1984) Regulation and information in a continuing relationship. *Inform. Econom. Policy* 1(3):267–302.
Battaglini M (2005) Long-term contracting with Markovian customers. *Amer. Econom. Rev.* 95(3):637–658.
Battaglini M (2007) Optimality and renegotiation in dynamic contracting. *Games Econom. Behav.* 60(2):213–246.
Battaglini M, Lamba R (2012) Optimal dynamic contracting. Working paper, Princeton University, Princeton, NJ.
Bergemann D, Said M (2011) Dynamic auctions: A survey. Cochran JJ, Cox LAJ, Keskinocak P, Kharoufeh JP, Smith JC, eds. *Wiley Encyclopedia of Operations Research and Management Science* (John Wiley & Sons, Hoboken, NJ).
Bergemann D, Välimäki J (2010) The dynamic pivot mechanism. *Econometrica* 78:771–789.

Boleslavsky R, Said M (2012) Progressive screening: Long-term contracting with a privately known stochastic process. *Rev. Econom. Stud.* 80(1):1–34.

Cavallo R, Parkes DC, Singh S (2007) Efficient online mechanisms for persistent, periodically inaccessible self-interested agents. Working paper, Harvard University, Cambridge, MA.

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

Deb R (2008) Optimal contracting of new experience goods. Working paper, University of Toronto, Toronto, Ontario.

Edelman B, Ostrovsky M, Schwarz M (2007) Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. *Amer. Econom. Rev.* 97(1):242–259.

Ëso P, Szentes B (2007) Optimal information disclosure in auctions and the handicap auction. *Rev. Econom. Stud.* 74(3):705–731.

Gallien J (2006) Dynamic mechanism design for online commerce. *Oper. Res.* 54(2):291–310.

Gershkov A, Moldovanu B (2009) Dynamic revenue maximization with heterogeneous objects: A mechanism design approach. *Amer. Econom. J.: Microeconomics* 1(2):168–198.

Gittins JC (1989) *Allocation Indices for Multi-Armed Bandits* (Wiley, London).

Mahdian M, Tomak K (2007) Pay-per-action model for online advertising. *Proc. 1st Internat. Workshop on Data Mining and Audience Intelligence for Advertising*, 1–6.

Milgrom P, Segal I (2002) Envelope theorems for arbitrary choice sets. *Econometrica* 70(2):583–601.

Myerson R (1981) Optimal auction design. *Math. Oper. Res.* 6(1):58–73.

Myerson R (1986) Multistage games with communications. *Econometrica* 54(2):323–358.

Nazerzadeh H, Saberi A, Vohra R (2013) Dynamic pay-per-action mechanisms and applications to online advertising. *Oper. Res.* 61(1): 98–111.

Pai M, Vohra R (2013) Optimal dynamic auctions and simple index rules. *Math. Oper. Res.*, ePub ahead of print May 6, http://dx.doi.org/10.1287/moor.2013.0595.

Parkes D (2007) Online mechanisms. Nisan N, Roughgarden T, Tardos E, Vazirani VV, eds. *Algorithmic Game Theory* (Cambridge University Press, Cambridge, UK), 441–442.

Pavan A, Segal I, Toikka J (2011) Dynamic mechanism design: Incentive compatibility, profit maximization and information disclosure. Working paper, Northwestern University, Evanston, IL.

Roberts K (1979) The characterization of implementable choice rules. Laffont, JJ, ed. *Aggregation and Revelation of Preferences* (Elsevier, Amsterdam), 321–349.

Said M (2012) Auctions with dynamic populations: Efficiency and revenue maximization. *J. Econom. Theory* 147(6):2419–2438.

Skrzypacz A, Board S (2010) Revenue management with forward-looking buyers. Working paper, Stanford University, Stanford, CA.

Vulcano G, van Ryzin G, Maglaras C (2002) Optimal dynamic auctions for revenue management. *Management Sci.* 48(11):1388–1407.

Whittle P (1982) *Optimization Over Time*, Vol. 1 (Wiley, Chichester, UK).

Zhang H (2012) Analysis of a dynamic adverse selection model with asymptotic efficiency. *Math. Oper. Res.* 37(3):450–474.

Sham M. Kakade is a senior research scientist at Microsoft Research, New England. His research focus is on designing scalable and efficient algorithms for machine learning and artificial intelligence. He had been an associate professor of statistics at the Wharton School at the University of Pennsylvania and an assistant professor at the Toyota Technological Institute.

Ilan Lobel is an assistant professor in information, operations, and management science at the Stern School of Business, New York University. He is interested in questions of pricing, learning, and contract design for online, dynamic, and networked markets.

Hamid Nazerzadeh is an assistant professor in the Information and Operations Management Department at the Marshall School of Business, University of Southern California. His research interests include market design, revenue management, and optimization algorithms. He holds several patents on Internet advertising and cloud computing services.

Optimal Dynamic Mechanism Design and the Virtual Pivot Mechanism

Sham Kakade Ilan Lobel Hamid Nazerzadeh

Microsoft Research, New York University and University of Southern California

skakade@microsoft.com, ilobel@stern.nyu.edu, hamidnz@marshall.usc.edu

Appendix

Appendix A: Proofs for Section 3

Lemma A.1 *For any reporting strategy $y \to z$ and initial type x, the partial derivative of the expected value of agent i $V_i^{y\to z}(x)$ (see definition in Eq. (10)) with respect to x exists and is:*

$$\frac{\partial V^{y\to z}(x)}{\partial x} = \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \frac{\partial}{\partial s_{i,0}} v_{i,t}(a^t, s_{i,0}, s_{i,1}, ..., s_{i,t}) \;\Big|_{s_{i,0}=x}\right]$$

(where the expectation is under $y \to z$ and $\mathcal{T}_{-i}$). Furthermore, it is bounded by

$$\left|\frac{\partial V_i^{y\to z}(x)}{\partial x}\right| \le \frac{\bar{V}}{1-\delta}.$$

Proof: From Assumption 2.2, we have that for all i, t, a, x and $s_{i,1}, ..., s_{i,t}$,

$$\left|\frac{\partial}{\partial x} v_{i,t}(a^t, x, s_{i,1}, ..., s_{i,t})\right| \le \bar{V} < \infty.$$

Therefore, by Lebesgue's Dominated Convergence Theorem, the partial derivative $\frac{\partial \bar{V}^{y\to z}(x)}{\partial x}$ exists,

$$\frac{\partial V^{y\to z}(x)}{\partial x} = \frac{\partial}{\partial x}\mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t v_{i,t}(a^t, x, s_{i,1}, ..., s_{i,t})\right] = \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \frac{\partial}{\partial x} v_{i,t}(a^t, x, s_{i,1}, ..., s_{i,t})\right]$$

and $|\frac{\partial V^{y\to z}(x)}{\partial x}| \le \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \bar{V}\right] = \frac{\bar{V}}{1-\delta}$.

Proof of Lemma 3.1 Any strategy available to the agents in the relaxed environment is a feasible strategy in the dynamic environment. Therefore, if all other agents are truthful, any profitable deviation from the truthful strategy in the relaxed environment implies a profitable deviation in the dynamic environment. Since no such profitable deviations exist in the dynamic environment, we obtain that the mechanism $\mathcal{M}$ is incentive compatible in the relaxed environment. Therefore, the optimal revenue in the relaxed environment provides an upper bound on the revenue in the dynamic environment.

Proof of Lemma 3.2 For consistency with the notation used in the rest of the paper, we represent the utility of agent i with initial type $s_{i,0} = z'$ and reporting his initial type as $\hat{s}_{i,0} = z$ by $U_i^{z\to z}(z')$, assuming all other agents are truthful. Respectively, $V_i^{z\to z}(z')$ and $P_i^{z\to z}(z')$ represent the expected discounted value and payment of agent i under initial type z' and reported initial type z'.

The expected utility of agent i under reporting strategy $z \to z$ and initial type x is

$$U_i^{z\to z}(z) = V_i^{z\to z}(z) - P_i^{z\to z}(z). \tag{27}$$

Under the same reporting strategy $z \to z$, but under initial type z', the utility of agent i is

$$U_i^{z\to z}(z') = V_i^{z\to z}(z') - P_i^{z\to z}(z'). \tag{28}$$

The payments are functions only of reported types, not true types, and therefore, $P_i^{z\to z}(z) = P_i^{z\to z}(z')$. Therefore, for any $z \neq z'$, combining Eqs. (27) and (28) yields

$$\frac{U_i^{z\to z}(z) - U_i^{z\to z}(z')}{z - z'} = \frac{V_i^{z\to z}(z) - V_i^{z\to z}(z')}{z - z'}.$$

At the same time, if $z > z'$, incentive compatibility yields $U_i^{z'\to z'}(z') \geq U_i^{z\to z}(z')$, hence

$$\frac{U_i^{z\to z}(z) - U_i^{z'\to z'}(z')}{z - z'} \leq \frac{U_i^{z\to z}(z) - U_i^{z\to z}(z')}{z - z'}.$$

Since the partial derivative $\frac{\partial V^{z\to z}(x)}{\partial x}$ exists for all x (see Lemma A.1), we can take the limit as $z' \uparrow z$ and obtain that the left-hand side derivative of $U_i^{z\to z}(z)$ satisfies

$$\frac{d_- U_i^{z\to z}(z)}{dz} \leq \frac{\partial V_i^{z\to z}(s)}{\partial s}\Big|_{s=z}.$$

Using the same argument for $z' > z$, we obtain that the right-hand side derivative of $U_i^{z\to z}(z)$ satisfies

$$\frac{d_+ U_i^{z\to z}(z)}{dz} \geq \frac{\partial V_i^{z\to z}(s)}{\partial s}\Big|_{s=z}.$$

Since $|\frac{\partial V_i^{z\to z}(s)}{\partial s}|$ is bounded by $\frac{\bar{V}}{1-\delta}$ by Lemma A.1, we get that the absolute value of both the left-hand and right-hand side derivatives of $U_i^{z\to z}(z)$ are also bounded by $\frac{\bar{V}}{1-\delta}$. The function $U_i^{z\to z}(z)$ is, therefore, $\frac{\bar{V}}{1-\delta}$-Lipschitz-continuous and, thus, differentiable almost everywhere. At all points where the derivative exists, $\frac{dU_i^{z\to z}(z)}{dz} = \frac{\partial V_i^{z\to z}(s)}{\partial s}\Big|_{s=z}$. Therefore, the envelope condition follows:

$$U_i^{x\to x}(x) - U_i^{x'\to x'}(x') = \int_{x'}^{x} \frac{dU_i^{z\to z}(z)}{dz} dz = \int_{x'}^{x} \frac{\partial V_i^{z\to z}(s)}{\partial s}\Big|_{s=z} dz. \tag{29}$$

Plugging in the result from Lemma A.1, we obtain the desired result.

Proof of Lemma 3.3 For notational convenience, we write:

$$\frac{\partial v_{i,t}(a^t, s_{i,0}, s_{i,1}, ..., s_{i,t})}{\partial s_{i,0}} \; |_{s_{i,0}=s_{i,0}} = \frac{\partial v_{i,t}(a^t, s_i^t)}{\partial s_{i,0}}$$

where the s_i^t implicitly depends on the first signal.

Consider first the utility $U_i^{\mathcal{M}}(s)$ of an agent i under an initial type profile s, which is given by

$$U_i^{\mathcal{M}}(s_{i,0}, s_{-i,0}) - U_i^{\mathcal{M}}(0, s_{-i,0}) = \int_0^{s_{i,0}} \mathbb{E}\left[\sum_{t=1}^{\infty} \delta^t \frac{\partial v_{i,t}(a^t, s_i^t)}{\partial s_{i,0}} \middle| s_{i,0} = z, s_{-i,0}\right] dz.$$

from Lemma 3.2. Taking the expectation of this term over all possible first period signals $s_{1,0}, ..., s_{n,0}$, we obtain

$$\mathbb{E}[U_i^{\mathcal{M}}(s_{i,0}, s_{-i,0}) - U_i^{\mathcal{M}}(0, s_{-i,0})] = \int_0^1 \left(\int_0^{s_{i,0}} \mathbb{E}\left[\sum_{t=1}^{\infty} \delta^t \frac{\partial v_{i,t}(a^t, s_i^t)}{\partial s_{i,0}} \middle| s_{i,0} = z\right] dz\right) f_i(s_{i,0}) ds_{i,0}.$$

Inverting the order of integration,

$$\begin{aligned}\mathbb{E}[U_i^{\mathcal{M}}(s_{i,0},s_{-i,0})-U_i^{\mathcal{M}}(0,s_{-i,0})] &= \int_0^1\int_z^1 \mathbb{E}\left[\sum_{t=1}^{\infty}\delta^t \frac{\partial v_{i,t}(a^t,s_i^t)}{\partial s_{i,0}}\middle| s_{i,0}=z\right] f_i(s_{i,0})ds_{i,0}dz \\ &= \int_0^1 \mathbb{E}\left[\sum_{t=1}^{\infty}\delta^t \frac{\partial v_{i,t}(a^t,s_i^t)}{\partial s_{i,0}}\middle| s_{i,0}=z\right](1-F_i(z))dz.\end{aligned}$$

By multiplying and dividing the right-hand side of the equation above by the density $f_i(z)$ we obtain an unconditional expectation,

$$\mathbb{E}[U_i^{\mathcal{M}}(s_{i,0},s_{-i,0})-U_i^{\mathcal{M}}(0,s_{-i,0})] = \mathbb{E}\left[\sum_{t=1}^{\infty}\delta^t \frac{1-F_i(s_{i,0})}{f_i(s_{i,0})}\frac{\partial v_{i,t}(a^t,s_i^t)}{\partial s_{i,0}}\right].$$

Now note that the discounted sum of payments $\mathbb{E}[\sum_{t=1}^{\infty}\delta^t p_{i,t}]$ is equal to the expected discounted valuation of agent i – $\mathbb{E}[\sum_{t=1}^{\infty}\delta^t v_{i,t}(a^t,s_i^t)]$ – minus her utility, which yields the claim.

Proof of Lemma 3.4 Observe that for multiplicatively-separable value functions

$$\frac{\partial v_{i,t}(a^t,s_i^t)}{\partial s_{i,0}} = A_i'(s_{i,0})B_{i,t}(a^t,s_{i,1},...,s_{i,t})$$

and, therefore, Eqs. (7) and (6) are identical. Similarly, for additively-separable functions,

$$\frac{\partial v_{i,t}(a^t,s_i^t)}{\partial s_{i,0}} = A_i'(s_{i,0})C_i(a_t)$$

and, therefore, Eqs. (7) and (6) are again identical.

Proof of Corollary 3.1 For an IC mechanism $\mathcal{M}$, the expected discounted sum of payments by agent i is equal to

$$\mathbb{E}\left[\sum_{t=0}^{\infty}\delta^t p_{i,t}\right] = \mathbb{E}\left[\sum_{t=0}^{\infty}\delta^t\big(\alpha_i(s_{i,0})v_{i,t}(a^t,s_i^t)+\beta_{i,t}(a^t,s_{i,0})\big)\right] - \mathbb{E}\left[U_i^{\mathcal{M},\mathcal{T}}(s_{i,0}=0)\right]$$

by taking expectations over $s_{-i,0}$ (see Eq. (7)). Since the mechanism satisfies IR, $\mathbb{E}\left[U_i^{\mathcal{M},\mathcal{T}}(s_{i,0}=0)\right]\geq 0$ and, therefore,

$$\mathbb{E}\left[\sum_{t=0}^{\infty}\delta^t p_{i,t}\right] \leq \mathbb{E}\left[\sum_{t=0}^{\infty}\delta^t\big(\alpha_i(s_{i,0})v_{i,t}(a^t,s_i^t)+\beta_{i,t}(a^t,s_{i,0})\big)\right].$$

The profit of $\mathcal{M}$ is given by the sum of payments minus the cost of actions (see Eq. (4)),

$$\text{Profit}^{\mathcal{M}} \leq \mathbb{E}\left[\sum_{t=1}^{\infty}\delta^t\left(\sum_{i=1}^{n}\Big(\alpha_i(s_{i,0})v_{i,t}(a^t,s_i^t)+\beta_{i,t}(a^t,s_{i,0})\Big)-c_t(a^t)\right)\right].$$

The bound above is valid for all IC and IR mechanisms. By maximizing over the set of all possible allocation rules (payment rules do not enter the equation above), we obtain the desired result.

Proof of Lemma 3.5 The expected utility of agent i under reporting strategy $x' \to z$ and initial type z is

$$U_i^{x'\to z}(z) = V_i^{x'\to z}(z) - P_i^{x'\to z}(z), \tag{30}$$

where $P_i^{x'\to z}(z)$ is the expected discounted sum of payments of agent i under reporting strategy $x' \to z$ and initial type z (see similar definitions of $U_i^{x'\to z}(z)$ and $V_i^{x'\to z}(z)$ in Eqs. (9) and (10)). Under the same reporting strategy $x' \to z$, but under initial type z', the utility of agent i is

$$U_i^{x'\to z}(z') = V_i^{x'\to z}(z') - P_i^{x'\to z}(z'). \tag{31}$$

The payments are functions only of reported types, not true types, and therefore, $P_i^{x'\to z}(z) = P_i^{x'\to z}(z')$. Therefore, for any $z \neq z'$, combining Eqs. (30) and (31) yields

$$\frac{U_i^{x'\to z}(z) - U_i^{x'\to z}(z')}{z - z'} = \frac{V_i^{x'\to z}(z) - V_i^{x'\to z}(z')}{z - z'}.$$

Periodic ex-post IC guarantees that $U_i^{x'\to z'}(z') \geq U_i^{x'\to z}(z')$. Therefore, for any $z > z'$,

$$\frac{U_i^{x'\to z}(z) - U_i^{x'\to z'}(z')}{z - z'} \leq \frac{U_i^{x'\to z}(z) - U_i^{x'\to z}(z')}{z - z'}.$$

Since the partial derivative $\frac{\partial V^{z\to z}(x)}{\partial x}$ exists for all x (see Lemma A.1), we can take the limit as $z' \uparrow z$ and obtain that the left-hand side derivative of $U_i^{x'\to z}(z)$ for any constant x' satisfies

$$\frac{d_- U_i^{x'\to z}(z)}{dz} \leq \frac{\partial V_i^{x'\to z}(s)}{\partial s}\Big|_{s=z}.$$

Using the same argument for $z' > z$, we obtain that the right-hand side derivative of $U_i^{x'\to z}(z)$ satisfies

$$\frac{d_+ U_i^{x'\to z}(z)}{dz} \geq \frac{\partial V_i^{x'\to z}(s)}{\partial s}\Big|_{s=z}.$$

Since $|\frac{\partial V_i^{x'\to z}(s)}{\partial s}|$ is bounded by $\frac{\bar{V}}{1-\delta}$ by Lemma A.1, we get that the absolute value of both the left-hand and right-hand side derivatives of $U_i^{x'\to z}(z)$ are also bounded by $\frac{\bar{V}}{1-\delta}$. The function $U_i^{x'\to z}(z)$ is, therefore, $\frac{\bar{V}}{1-\delta}$-Lipschitz-continuous and, thus, differentiable almost everywhere. At all points where the derivative exists, $\frac{dU_i^{x'\to z}(z)}{dz} = \frac{\partial V_i^{x'\to z}(s)}{\partial s}\Big|_{s=z}$. Therefore, the envelope condition follows:

$$U_i^{x'\to x}(x) - U_i^{x'\to x'}(x') = \int_{x'}^{x} \frac{dU_i^{x'\to z}(z)}{dz} dz = \int_{x'}^{x} \frac{\partial V_i^{x'\to z}(s)}{\partial s}\Big|_{s=z} dz.$$

Proof of Lemma 3.6 The envelope condition from the relaxed environment (see Lemma 3.2) also applies to this setting since a deviation that is feasible in the relaxed environment (that is, using reporting strategy $z \to z$ for an initial type z') is also feasible in the dynamic environment. Therefore, if the mechanism is incentive compatible, then it satisfies Eq. (29), which is identical to Eq. (15).

To see that IC implies the dynamic monotonicity condition in Eq. (16), simply note that IC is equivalent to Eq. (11) and Eqs. (15) and (13) are respectively equal to the left-hand and the right-hand side of Eq. (11). We thus obtain that IC implies Eq. (16).

We now show that if both Eqs. (15) and (16) hold, then the mechanism is IC. If both equations hold, then for all x and x',

$$U_i^{x\to x}(x) - U_i^{x'\to x'}(x') = \int_{x'}^{x} \frac{\partial V_i^{z\to z}(s)}{\partial s}\Big|_{s=z} dz \geq \int_{x'}^{x} \frac{\partial V_i^{x'\to z}(s)}{\partial s}\Big|_{s=z} dz = U_i^{x'\to x}(x) - U_i^{x'\to x'}(x'),$$

where the last equality follows from Lemma 3.5. The equation above is equivalent to IC (see Eq. (11)), when the mechanism is periodic ex-post IC for $t \geq 1$.

Appendix B: Proofs for Section 4

Lemma B.1 *Suppose Assumptions 4.1 and 4.2 hold. Then α_i is strictly increasing for multiplicatively separable functions and $\beta_{i,t}$ is strictly increasing for additively separable functions.*

Proof: For simplicity of nation, let $s = s_{i,0}$. Also, let $\eta_i(s)$ denote the hazard rate, i.e.,

$$\eta_i(s) = \frac{f_i(s)}{1 - F_i(s)}.$$

In the additive case,

$$\frac{\partial \beta_{i,t}(a^t, s)}{\partial s} = \frac{\eta_i'(s)}{\eta_i^2(s)} A_i'(s) C_{i,t}(a^t) - \frac{1}{\eta_i(s)} A_i''(s) C_{i,t}(a^t)$$

where $(\cdot)'$ denotes a partial derivative with respect to s. By the assumptions that A_i is concave and strictly increasing, and the hazard rate is positive and strictly increasing, we have that the above has the same sign as $C_{i,t}$. In the multiplicative case, first note that $\alpha_i(s) = 1 - \frac{1}{\eta_i(s)}(\log A_i(s))'$. Therefore,

$$\alpha_i'(s) = \frac{\eta_i'(s)}{\eta_i^2} \frac{A_i'(s)}{A_i(s)} - \frac{1}{\eta_i(s)} (\log A_i(s))''$$

which is positive by the assumption.

Proof of Lemma 4.1 If agents are truthful, by Eq. (24), the expected payment of each agent i given $s_{i,0}$ is equal to $\max\{p_i^\star(s_{i,0}), 0\}$, where 0 occurs if agent i is excluded from the system ($i \notin a_i$). Namely,

$$p_i^\star(\hat{s}_0) = V(s_i^t) - \int_0^{\hat{s}_{i,0}} \frac{\partial V_i^{z \to z}(s_{i,0}, \hat{s}_{-i,0})}{\partial s_{i,0}} \Big|_{s_{i,0}=z} dz \tag{32}$$

where

$$\frac{\partial V_i^{z \to z}(s_{i,0}, \hat{s}_{-i,0})}{\partial s_{i,0}} \Big|_{s_{i,0}=z} = \mathbb{E}\left[\sum_{t=1}^{\infty} \delta^t \frac{\partial v_{i,t}(q^{\star t}, s_{i,0}, s_{i,1}, \dots s_{i,t})}{\partial s_{i,0}} \Big|_{s_{i,0}=z} \middle| s_{i,0} = z, s_{-i,0} = \hat{s}_{-i,0} \right]$$

For notational convenience, we write:

$$\frac{\partial v_{i,t}(a^t, s_{i,0}, s_{i,1}, \dots, s_{i,t})}{\partial s_{i,0}} \Big|_{s_{i,0}=s_{i,0}} = \frac{\partial v_{i,t}(a^t, s_i^t)}{\partial s_{i,0}}$$

where the s_i^t implicitly depends on the first signal. The expected payment of agent i is equal to:

$$\int_0^1 \max\{p_i^\star(s, s_{0,-i}), 0\} f_i(s) ds$$
$$= \int_0^1 \left(\mathbb{E}\left[\sum_{t=1}^{\infty} \delta^t v_{i,t}(q^{\star t}, s_i^t) \middle| s_{i,0} = s, s_{-i,0} \right] - \int_0^s \mathbb{E}\left[\sum_{t=1}^{\infty} \delta^t \frac{\partial v_{i,t}(q^{\star t}, s_i^t)}{\partial s_{i,0}} \middle| s_{i,0} = z, s_{-i,0} \right] dz \right) f_i(s) ds,$$

where we can drop the max with zero since the agent obtains value zero at all periods when she is excluded from the system. By changing the order of integration, we have

$$\int_0^1 \max\{p_i^\star(s, s_{0,-i}), 0\} f_i(s) ds$$

$$
\begin{aligned}
&= \int_0^1 \left(\mathbb{E}\left[\sum_{t=1}^{\infty} \delta^t \left(v_{i,t}(q^{\star t}, s_i^t) - \frac{1-F_i(s)}{f_i(s)} \frac{\partial v_{i,t}(q^{\star t}, s_i^t)}{\partial s_{i,0}} \right) \middle| s_{i,0} = s, s_{-i,0} \right] \right) f_i(s) ds \\
&= \int_0^1 \left(\mathbb{E}\left[\sum_{t=1}^{\infty} \delta^t \left(\alpha_i(s_{i,0}) v_{i,t}(q^{\star t}, s_i^t) + \beta_{i,t}(q^{\star t}, s_{i,0}) \right) \middle| s_{i,0} = s, s_{-i,0} \right] \right) f_i(s) ds \\
&= \mathbb{E}\left[\sum_{t=1}^{\infty} \delta^t \left(\alpha_i(s_{i,0}) v_{i,t}(q^{\star t}, s_i^t) + \beta_{i,t}(q^{\star t}, s_{i,0}) \right) \middle| s_{-i,0} \right] \qquad (33)
\end{aligned}
$$

Therefore, the profit of the mechanism matches the upper-bound provided in Corollary 3.1. Hence, to prove the optimality, it suffices to show that the mechanism is individually rational. By construction, we have the utility of agent i equal to 0 if $s_{i,0} = 0$ for any $s_{-i,0}$. Therefore,

$$
U_i(s_0) = \int_0^{s_{i,0}} \mathbb{E}\left[\sum_{t=1}^{\infty} \delta^t \frac{\partial v_{i,t}(q^{\star t}, s_{i,0}, s_{i,1}, \dots s_{i,t})}{\partial s_{i,0}} |_{s_{i,0}=z} \middle| s_{i,0} = z, s_{-i,0} \right] dz.
$$

By Assumption 4.2, $\frac{\partial v_{i,t}(q^{\star t}, s_{i,0}, s_{i,1}, \dots s_{i,t})}{\partial s_{i,0}}$ is non-negative. Hence, the mechanism is individually rational. Precisely, periodic ex-post IR at time 0.

Proof of Lemma 4.2 Define $u_{i,t}$ to be the instantaneous utility of agent i at time t. We get

$$
\begin{aligned}
u_{i,t} &= v_{i,t}(a^{\star t}, s_i^t) - p_{i,t} \\
&= \left(v_{i,t}(a^{\star t}, s_i^t) - v_{i,t}(a^{\star t}, \hat{s}_i^t) \right) + \frac{m_{i,t}}{\hat{\alpha}_i} \\
&= v_{i,t}(a^{\star t}, s_i^t) + \frac{\hat{\beta}_i(a^{\star t})}{\hat{\alpha}_i} \\
&\quad + \frac{1}{\hat{\alpha}_i} \left(\sum_{j \neq i} \left(\hat{\alpha}_j v_{j,t}(a^{\star t}, \hat{s}^{j,t}) \right) - c_t(a^{\star t}) - W_{-i}^{(\hat{\alpha},\hat{\beta})}(a^{\star t-1}, \hat{s}^t) + \delta \mathbb{E}\left[W_{-i}^{(\hat{\alpha},\hat{\beta})}(a^{\star t}_{-i}, \hat{s}^{t+1}) \right] \right)
\end{aligned}
$$

The last equality follows from Eq. (21). We dropped the conditioning of $W_{-i}^{(\hat{\alpha},\hat{\beta})}(a^{\star t}, \hat{s}^{t+1})$ on $s^t = \hat{s}_t$, $a^{\star t}$, and $a^{\star}_{-i,t}$, as it is clear from the context. For ease of notation, let $s = s_0$. Because all agents except i are truthful, we have

$$
\begin{aligned}
u_{it} = \frac{1}{\hat{\alpha}_i} \Bigg(&\sum_{j=1}^{n} \left(\hat{\alpha}_j v_{j,t}(a^{\star t}, s_j^t) + \hat{\beta}_j(a^{\star t}) \right) - c_t(a^{\star t}) \\
&- W_{-i}^{(\alpha(s),\beta(s))}(a^{\star t-1}, s^t) + \delta \mathbb{E}\left[W_{-i}^{(\alpha(s),\beta(s))}(a^{\star t}_{-i}, s^{t+1}) \right] \Bigg)
\end{aligned}
$$

If agent i is truthful and other agents are truthful, we have

$$
\sum_{t'=t}^{\infty} \delta^t u_{it'} = \frac{1}{\hat{\alpha}_i} \left(W^{(\alpha(s),\beta(s))}(a^{\star t-1}, s^t) - W_{-i}^{(\alpha(s),\beta(s))}(a^{\star t-1}, s^t) \right)
$$

Hence, the allocation rule is aligned with the incentive of agent i. She can maximize her utility by reporting truthfully.

Observe that agents with $\hat{\alpha}_i \leq 0$ would have been excluded. Hence, we have $\sum_{t'=t}^{\infty} u_{it'} \geq 0$. Therefore, the mechanism is periodic ex-post IR.

Proof of Lemma 4.3 Observe that Eq. (15) is followed from Lemma 4.1 and Eq. (33). To establish Eq. (16), we show that the inequality holds point-wise, i.e., if $x \geq x'$, then

$$\frac{\partial V_i^{x_i \to x_i}(s)}{\partial s}\Big|_{s=x_i} \geq \frac{\partial V_i^{x_i' \to x_i}(s)}{\partial s}\Big|_{s=x_i} \tag{34}$$

By Eq. (14), this is equivalent to

$$\mathbb{E}_{x_i \to x_i}\left[\sum_{t=0}^{\infty} \delta^t \frac{\partial v_{i,t}(a^{\star t}, s_i^t)}{\partial s_{i,0}} \ |_{s_{i,0}=s} \middle| s_{i,0} = x_i\right] \geq \mathbb{E}_{x_i' \to x_i}\left[\sum_{t=0}^{\infty} \delta^t \frac{\partial v_{i,t}(a'^t, s_i^t)}{\partial s_{i,0}} \ |_{s_{i,0}=s} \middle| s_{i,0} = x_i\right] \tag{35}$$

where $\mathbb{E}_{x_i \to x_i}$ is the expectation under the stochastic process determined by agent i reporting according to $x_i \to x_i$ (while other agents are truthful) and $a^{\star t}$ represents the allocation at time t in this case. Similarly, for reporting strategy $x_i' \to x_i$, we use the notation $\mathbb{E}_{x_i' \to x_i}$ and represent the allocation at time t by $a^{'t}$.

Recall that we have:

$$v_{i,t}(a^t, s_i^t) - \frac{1 - F_i(s_{i,0})}{f_i(s_{i,0})} \frac{\partial v_{i,t}(a^t, s_i^t)}{\partial s_{i,0}} = \alpha_i(s_{i,0}) v_{i,t}(a^t, s_i^t) + \beta_{i,t}(a^t, s_{i,0})$$

Hence, we get

$$\frac{\partial v_{i,t}(a^t, s_i^t)}{\partial s_{i,0}} = \frac{f_i(s_{i,0})}{1 - F_i(s_{i,0})} \big((1 - \alpha_i(s_{i,0})) v_{i,t}(a^t, s_i^t) - \beta_{i,t}(a^t, s_i^t)\big) \tag{36}$$

Therefore, by Eq. (36), the inequality below is equivalent to the desired equation, Eq. (34):

$$\begin{aligned} &\mathbb{E}_{x_i \to x_i}\left[\sum_{t=1}^{\infty} \delta^t \left((1 - \alpha_i(x_i)) v_{i,t}(a^t, s_i^t) - \beta_{i,t}(a^t, x_i)\right)\right] \\ &\geq \mathbb{E}_{x_i' \to x_i}\left[\sum_{t=1}^{\infty} \delta^t \Big((1 - \alpha_i(x_i)) v_{i,t}(a'^t, s_i^t) - \beta_{i,t}(a'^t, x_i)\Big)\right] \end{aligned} \tag{37}$$

In the following we prove the inequality above. For $k \neq i$, define x_k and x_k' to be equal $s_{k,0}$. Because $a^\star$ and a' are optimal allocation rules with respect to $(\alpha(x), \beta(x))$ and $(\alpha(x'), \beta(x'))$, we have:

$$\begin{aligned} &\mathbb{E}_{x_i \to x_i}\left[\sum_{t=1}^{\infty} \delta^t \left(\sum_{j=1}^{n} \big(\alpha_j(x_j) v_{j,t}(a^{\star t}, s_j^t) + \beta_{j,t}(a^{\star t}, x_j)\big) - c_t(a^{\star t})\right)\right] \\ \geq\ &\mathbb{E}_{x_i' \to x_i}\left[\sum_{t=1}^{\infty} \delta^t \left(\sum_{j=1}^{n} \big(\alpha_j(x_j) v_{j,t}(a'^t, s_j^t) + \beta_{j,t}(a'^t, x_j)\big) - c_t(a'^t)\right)\right] \end{aligned}$$

and similarly

$$\begin{aligned} &\mathbb{E}_{x_i \to x_i}\left[\sum_{t=1}^{\infty} \delta^t \left(\sum_{j=1}^{n} \big(\alpha_j(x_j') v_{j,t}(a^{\star t}, s_j^t) + \beta_{j,t}(a^{\star t}, x_j)\big) - c_t(a^{\star t})\right)\right] \\ \leq\ &\mathbb{E}_{x_i' \to x_i}\left[\sum_{t=1}^{\infty} \delta^t \left(\sum_{j=1}^{n} \big(\alpha_j(x_j') v_{j,t}(a'^t, s_j^t) + \beta_{j,t}(a'^t, x_j')\big) - c_t(a'^t)\right)\right] \end{aligned}$$

Subtracting these inequalities we get:

$$\mathbb{E}_{x_i\to x_i}\left[\sum_{t=1}^{\infty}\delta^t\sum_{j=1}^{n}\Big((\alpha_j(x_j)-\alpha_j(x'_j))v_{j,t}(a^{\star t},s_j^t)+(\beta_{j,t}(a^{\star t},x_j)-\beta_{j,t}(a^{\star t},x'_j))\Big)\right]$$
$$\geq \mathbb{E}_{x'_i\to x_i}\left[\sum_{t=1}^{\infty}\delta^t\sum_{j=1}^{n}\Big((\alpha_j(x_j)-\alpha_j(x'_j))v_{j,t}(a'^t,s_j^t)+(\beta_{j,t}(a'^t,x_j)-\beta_{j,t}(a'^t,x'_j))\Big)\right]$$

Because for $k\neq i$, agents are truthful and $x'_k=x_k$, we have

$$\mathbb{E}_{x_i\to x_i}\left[\sum_{t=1}^{\infty}\delta^t\Big((\alpha_i(x_i)-\alpha_i(x'_i))v_{i,t}(a^{\star t},s_i^t)+(\beta_{i,t}(a^{\star t},x_i)-\beta_{i,t}(a^{\star t},x'_i))\Big)\right] \tag{38}$$
$$\geq \mathbb{E}_{x'_i\to x_i}\left[\sum_{t=1}^{\infty}\delta^t\Big((\alpha_i(x_i)-\alpha_i(x'_i))v_{i,t}(a'^t,s_i^t)+(\beta_{i,t}(a'^t,x_i)-\beta_{i,t}(a'^t,x'_i))\Big)\right]$$

Now suppose v_i is multiplicative separable (i.e., $\beta_{i,t}(\cdot,\cdot)=0$) and Assumption 4.2 holds — we consider the additive valuations later. Because $x\geq x'$, by Assumption 4.2 and Lemma B.1, we have $\alpha_i(x_i)>\alpha_i(x'_i)$; moreover $\alpha_i(x_i)$ is less than 1 for $x\in[0,1)$. Multiplying both sides of the inequality above by $\frac{1-\alpha_i(x_i)}{\alpha_i(x_i)-\alpha_i(x'_i)}$, yields the following:

$$\mathbb{E}_{x_i\to x_i}\left[\sum_{t=1}^{\infty}\delta^t(1-\alpha_i(x_i))v_{i,t}(a^{\star t},s_i^t)\right]\geq \mathbb{E}_{x'_i\to x_i}\left[\sum_{t=1}^{\infty}\delta^t(1-\alpha_i(x_i))v_{i,t}(a'^t,s_i^t)\right]$$

which is equivalent to Eq. (37) for multiplicative-separable valuations.

Now consider the case of additive-separable value functions. We have $\alpha_i(x)=\alpha_i(x')=1$. Plugging into Eq. (38) we get

$$\mathbb{E}_{x_i\to x_i}\left[\sum_{t=1}^{\infty}\delta^t(\beta_{i,t}(a^{\star t},x_i)-\beta_{i,t}(a^{\star t},x'_i))\right]\geq \mathbb{E}_{x'_i\to x_i}\left[\sum_{t=1}^{\infty}\delta^t(\beta_{i,t}(a'^t,x_i)-\beta_{i,t}(a'^t,x'_i))\right]$$

Recall that $\beta_{i,t}(a^t,x_i)=-\frac{1-F_i(x_i)}{f_i(x_i)}A'_i(x_i)C_{i,t}(a^t)$. Because $x\geq x'$, by Assumption 4.2 and Lemma B.1, we have $-\frac{1-F_i(x_i)}{f_i(x_i)}A'_i(x_i)>-\frac{1-F_i(x'_i)}{f_i(x'_i)}A'_i(x'_i)$. By multiplying both sides of the inequality above by $\frac{\frac{1-F_i(x_i)}{f_i(x_i)}A'_i(x_i)}{-\frac{1-F_i(x_i)}{f_i(x_i)}A'_i(x_i)+\frac{1-F_i(x'_i)}{f_i(x'_i)}A'_i(x'_i)}$, we get:

$$-\mathbb{E}_{x_i\to x_i}\left[\sum_{t=1}^{\infty}\delta^t\beta_{i,t}(a^{\star t},x_i)\right]\geq -\mathbb{E}_{x'_i\to x_i}\left[\sum_{t=1}^{\infty}\delta^t\beta_{i,t}(a'^t,x'_i)\right]$$

which produces Eq. (37) and, thus, completes the proof.

Appendix C: Proof for the Single Agent Case

Proof of Corollary 5.1 Simply note that under the VIRTUAL-PIVOT Mechanism, if the agent is allocated the item at any time t, the price she pays, under the VIRTUAL-PIVOT Mechanism, is not a function of her report at time t (or any report after $t=0$). Furthermore, the prices that the agent is charged at $t\geq 1$ are identical to that in the VIRTUAL-PIVOT Mechanism (see Eq. (22)). Also, the prices charged at $t=0$ is identical to that in the VIRTUAL-PIVOT Mechanism by construction.