

## MIT Open Access Articles

### *Dynamic Mechanism Design: A Myersonian Approach*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Pavan, Alessandro, Ilya Segal, and Juuso Toikka. "Dynamic Mechanism Design: A Myersonian Approach." *Econometrica* 82, no. 2 (2014): 601–653.

**As Published:** <http://dx.doi.org/10.3982/ECTA10269>

**Publisher:** The Econometric Society

**Persistent URL:** <http://hdl.handle.net/1721.1/96767>

**Version:** Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

**Terms of use:** Creative Commons Attribution-Noncommercial-Share Alike



# Dynamic Mechanism Design: A Myersonian Approach\*

Alessandro Pavan

Northwestern University

Ilya Segal

Stanford University

Juuso Toikka

Massachusetts Institute of Technology

December 8, 2013

## Abstract

We study mechanism design in dynamic quasilinear environments where private information arrives over time and decisions are made over multiple periods. Our first main result is a necessary condition for incentive compatibility that takes the form of an envelope formula for the derivative of an agent’s equilibrium expected payoff with respect to his current type. It combines the familiar marginal effect of types on payoffs with novel marginal effects of the current type on future ones that are captured by “impulse response functions.” The formula yields an expression for dynamic virtual surplus which is instrumental to the design of optimal mechanisms, and to the study of distortions under such mechanisms. Our second main result gives transfers that satisfy the envelope formula, and establishes a sense in which they are pinned down by the allocation rule (“revenue equivalence”). Our third main result is a characterization of PBE-implementable allocation rules in Markov environments, which yields tractable sufficient conditions that facilitate novel applications. We illustrate the results by applying them to the design of optimal mechanisms for the sale of experience goods (“bandit auctions”).

*JEL Classification Numbers:* D82, C73, L1.

*Keywords:* asymmetric information, stochastic processes, incentives, mechanism design, envelope theorems

---

\*We thank participants at various conferences, seminars, and institutions where this paper was presented. A special thank is to Thomas Chaney, Daniel Garrett, Li Hao, Narayana Kocherlakota, Andy Skrzypacz, Xianwen Shi, the editor Stephen Morris, and anonymous referees, for useful suggestions that helped us improve the paper. Pavan also thanks the hospitality of Collegio Carlo Alberto and the University of California Berkeley where part of this work was completed. This paper supersedes previous working papers “Revenue Equivalence, Profit Maximization and Transparency in Dynamic Mechanisms” by Segal and Toikka, “Long-Term Contracting in a Changing World” by Pavan, and “Infinite-Horizon Mechanism Design” by Pavan, Segal, and Toikka.

# 1 Introduction

We consider the design of incentive compatible mechanisms in a dynamic environment in which agents receive private information over time and decisions are made in multiple periods over an arbitrary time horizon. The model allows for serial correlation of the agents' information and for the dependence of this information on past allocations. For example, it covers as special cases problems such as allocation of private or public goods to agents whose valuations evolve stochastically over time, procedures for selling experience goods to consumers who refine their valuations upon consumption, and multi-period procurement under learning-by-doing. Because of the arbitrary time horizon, the model also accommodates problems where the timing of decisions is a choice variable such as when auctioning off rights for the extraction of a natural resource.

Our main results, Theorems 1–3, provide characterizations of dynamic local and global incentive-compatibility constraints that extend the Myersonian approach to mechanism design with continuous types (Myerson, 1981) to dynamic environments. We then apply these results to the design of optimal dynamic mechanisms. We focus on quasilinear environments where the agents' new private information is unidimensional in each period.<sup>1</sup> In order to rule out the possibility of full surplus extraction à la Cremer and McLean (1988), we assume throughout that this information is independent across agents conditional on the allocations observed by them. In addition to the methodological contribution, our results provide some novel concepts that facilitate a unified view of the existing literature and help to explain what drives distortions in optimal dynamic contracts.

The cornerstone of our analysis is a dynamic envelope theorem, Theorem 1, which, under appropriate regularity conditions, yields a formula for the derivative of an agent's expected equilibrium payoff with respect to his current private information, or type, in any perfect Bayesian incentive compatible mechanism.<sup>2</sup> Similarly to Mirrlees' (1971) first-order approach for static environments, this formula characterizes local incentive-compatibility constraints. It captures the usual direct effect of a change in the current type on the agent's utility as well as a novel indirect effect due to the induced change in the distribution of the agent's future types. The stochastic component of the latter is summarized by *impulse response functions* that describe how a change in the agent's current type propagates through his type process. Theorem 1 thus identifies the impulse response as the notion of stochastic dependence relevant for mechanism design. Our definition of the impulse response functions and the proof of Theorem 1 make use of the fact that any stochastic process can be constructed from a sequence of independent random variables. This observation was first used in

---

<sup>1</sup>By reinterpreting monetary payments as “utility from monetary payments,” all of our results on incentive compatibility trivially extend to non-quasilinear environments where the agents' utility from monetary payments (or, more generally, from some other instrument available to the designer) is independent of their private information and additively separable from their allocation utility. For example, this covers models typically considered in new dynamic public finance or in the managerial compensation literature.

<sup>2</sup>This envelope theorem may be useful also in other stochastic dynamic programming problems.

the context of mechanism design by Esó and Szentes (2007).

The envelope formula of Theorem 1 is independent of transfers, and thus applying it to the initial period yields a dynamic payoff equivalence result that generalizes the revenue equivalence theorem of Myerson (1981). On the other hand, given any dynamic allocation rule, the envelope formula can be used to construct payments, which satisfy local incentive compatibility constraints at all truthful histories. We show this in Theorem 2, which also provides a sense in which these payments are unique. In particular, in the single-agent case, the net present value of payments for any realized sequence of types for the agent is determined by the allocation rule up to a single scalar. This *ex-post payoff equivalence* extends to multiple agents under an additional condition (which, for instance, is satisfied if the evolution of types is independent of allocations) to pin down the expected net present value of payments conditional on the agent’s own type sequence, where the expectation is over the other agents’ type sequences.<sup>3</sup>

We then focus on Markov environments in order to characterize global incentive-compatibility constraints. We show in Theorem 3 that an allocation rule is implementable in a perfect Bayesian equilibrium if, and only if, it satisfies *integral monotonicity*. The Markov restriction implies that when this is the case, the allocation rule can be implemented, using payments from Theorem 2, in a *strongly truthful* equilibrium where the agents report truthfully on and off the equilibrium path. This allows us to restrict attention to one-shot deviations from truth-telling, and is the reason for our focus on Markov environments. It is instructive to note, however, that even if an agent’s current type is unidimensional, his report can affect allocations in multiple periods. Thus the static analog of our problem is one with unidimensional types but multi-dimensional allocations, which explains why the integral-monotonicity condition cannot be simplified without losing necessity.<sup>4</sup>

Theorem 3 facilitates formulating sufficient conditions for implementability that are stronger than necessary, but easier to verify. A special case is the notion of *strong monotonicity* typically considered in the literature, which requires that each agent’s allocation be increasing in his current and past reports in every period, and which is applicable to models where payoffs satisfy a single-crossing property and where type transitions are independent of allocations and increasing in past types in the sense of first-order stochastic dominance. Having identified the underlying integral-monotonicity condition, we are able to relax both the notion of monotonicity and the requirements on the environment. Heuristically, this amounts to requiring monotonicity only “on average” across time (*ex-post monotonicity*) or, weaker still, across time and future types (*average monotonicity*). We use these new conditions to establish the implementability of the optimal allocation rule in settings where strong monotonicity fails.

---

<sup>3</sup>The result is useful in the non-quasilinear models of footnote 1. There it determines the ex post net present value of the agent’s utility from monetary payments, and facilitates computing the cost-minimizing timing of payments.

<sup>4</sup>Implementability has been characterized in static models in terms of analogous conditions by Rochet (1987), and more recently by Carbajal and Ely (2013) and Berger, Müller, and Naeemi (2010).

The leading application for our results is the design of optimal mechanisms in Markov environments.<sup>5</sup> We adopt the first-order approach familiar from static settings where an allocation rule is found by solving a relaxed problem that only imposes local incentive-compatibility constraints and the lowest initial types’ participation constraints, and where a monotonicity condition is used to verify the implementability of the rule. The envelope formula from Theorem 1 can be used as in static settings to show that the principal’s problem is then to maximize expected virtual surplus, which is only a function of the allocation rule. This is a Markov decision problem, and hence it can be solved using standard methods. We then use integral monotonicity from Theorem 3 to verify that the solution is implementable, possibly by checking one of the sufficient conditions discussed above. When this is the case, the optimal payments can be found by using Theorem 2. If for each agent the lowest initial type is the one worst off under the candidate allocation rule (which is the case, for example, when utilities are increasing in own types and transitions satisfy first-order stochastic dominance), then the participation constraints of all initial types are satisfied, and the mechanism so constructed is an optimal dynamic mechanism.

The impulse response functions play a central role in explaining the direction and dynamics of distortions in optimal dynamic mechanisms. As in static settings, distortions are introduced to reduce the agents’ expected information rents, as computed at the time of contracting. However, because of the serial correlation of types, it is optimal to distort allocations not only in the initial period, but at every history at which the agent’s type is responsive to his initial type, as measured by the impulse response function. We illustrate by means of a buyer-seller example that this can lead to the distortions being non-monotone in the agent’s reports and over time. The optimal allocation rule in the example is not strongly monotone, and hence the new sufficiency conditions derived from integral monotonicity are instrumental for uncovering these novel dynamics.

Similarly to static settings, the first-order approach outlined above yields an implementable allocation rule only under fairly stringent conditions, which are by no means generic. We provide some sufficient conditions on the primitives that guarantee that the relaxed problem has a solution that satisfies strong monotonicity, but as is evident from above, such conditions are far from being necessary. We illustrate the broader applicability of the tools by solving for optimal “bandit auctions” of experiment goods in a setting where bidders update their values upon consumption. The optimal allocation there violates strong monotonicity, but satisfies average monotonicity.

We conclude the Introduction by commenting on the related literature. The rest of the paper is then organized as follows. We describe the dynamic environment in Section 2, and present our results on incentive compatibility and implementability in Section 3. We then apply these results to the design of optimal dynamic mechanisms in Section 4, illustrating the general approach by deriving

---

<sup>5</sup>In the supplementary material, we discuss how our characterization of incentive compatibility for Markov environments can be used to derive sufficient conditions for implementability in some classes of non-Markov environments, and to extend our results on optimal mechanisms to such environments.

the optimal bandit auction in Section 5. We conclude in Section 6. All proofs omitted in the main text are in the Appendix. Additional results can be found in the supplementary material.

## 1.1 Related Literature

The literature on optimal dynamic mechanism design goes back to the pioneering work of Baron and Besanko (1984), who used the first-order approach in a two-period single-agent setting to derive an optimal mechanism for regulating a natural monopoly. They characterized optimal distortions using an “informativeness measure,” which is a two-period version of our impulse response function. More recently, Courty and Li (2000) considered a similar model to study optimal advanced ticket sales, and also provided some sufficient conditions for a dynamic allocation rule to be implementable. Esó and Szentes (2007) then extended the analysis to multiple agents in their study of optimal information revelation in auctions.<sup>6</sup> They orthogonalized an agent’s future information by generating the randomness in his second-period type through an independent shock, which corresponds to a two-period version of our state representation. We build on some of the ideas and results in these papers, and special cases of some of our results can be found in them. We comment on the connections at the relevant points in the analysis. In particular, we discuss the role of the state representation in the Concluding Remarks (Section 6) having first presented our results.

Whereas the aforementioned works considered two-period models, Besanko (1985) and Battaglini (2005) characterized the optimal infinite-horizon mechanism for an agent whose type follows a Markov process, with Besanko considering a linear AR(1) process over a continuum of states, and Battaglini a two-state Markov chain. Their results were qualitatively different: Besanko (1985) found the allocation in each period to depend only on the agent’s initial and current type, and to be distorted downward at each finite history with probability 1. In contrast, Battaglini (2005) found that once the agent’s type turns high, he consumes at the efficient level irrespective of his subsequent types. Our analysis shows that the relevant property of the type processes that explains these findings, and the dynamics of distortions more generally, is the impulse response of future types to a change in the agent’s initial private information.<sup>7</sup>

Optimal mechanisms in a multi-agent environment with an infinite horizon were first considered by Board (2007). He extended the analysis of Esó and Szentes (2007) to a setting where the timing of the allocation is endogenous, so that the principal is selling options. Subsequent to the first version of our manuscript, Kakade et al (2011) considered a class of allocation problems that generalize Board’s model as well as our bandit auctions, and showed that the optimal mechanism is a virtual version of the dynamic pivot mechanism of Bergemann and Välimäki (2010). We comment on the

---

<sup>6</sup>See Riordan and Sappington (1987) for an early contribution with many agents.

<sup>7</sup>Battaglini’s (2005) model with binary types is not formally covered by our analysis. However, we discuss in the supplementary material how impulse responses can be adapted to discrete type models.

connection to Kakade et al in Section 5 after presenting our bandit auction.

That the literature on optimal dynamic mechanisms has focused on relatively specific settings reflects the need to arrive at a tractable optimization problem over implementable allocation rules. In contrast, when designing efficient (or expected surplus maximizing) mechanisms, the desired allocation rule is known a priori. Accordingly, dynamic generalizations of the static Vickrey-Clarke-Groves and expected-externality mechanisms were recently introduced by Bergemann and Välimäki (2010) and Athey and Segal (2013) for very general, quasilinear private-value environments.<sup>8</sup>

A growing literature considers both efficient and profit-maximizing dynamic mechanisms in settings where each agent receives only one piece of private information, but where the agents or objects arrive stochastically over time as in, for example, Gershkov and Moldovanu (2009). The characterization of incentive compatibility in such models is static, but interesting dynamics emerge from the optimal timing problem faced by the designer. We refer the reader to the excellent recent survey by Bergemann and Said (2011).<sup>9</sup>

Our work is also related to the literature on dynamic insurance and optimal taxation. While the early literature following Green (1987) and Atkeson and Lucas (1992) assumed i.i.d. types, the more recent literature has considered persistent private information (e.g., Fernandes and Phelan (2000), Kocherlakota (2005), Albanesi and Sleet (2006), or Kapicka (2013)). In terms of the methods, particularly related are Farhi and Werning (2013) and Golosov et al (2011), who used a first-order approach to characterize optimal dynamic tax codes. There is also a continuous-time literature on contracting with persistent private information that uses Brownian motion, in which impulse responses are constant, simplifying the analysis. See Williams (2011) and the references therein.

Our analysis of optimal mechanisms assumes that the principal can commit to the mechanism he is offering, and hence the dynamics are driven by changes in the agents' information. In contrast, the literature on dynamic contracting with adverse selection and limited commitment typically assumes constant types and generates dynamics through lack of commitment (see, for example, Laffont and Tirole (1988), or, for more recent work, Skreta (2006) and the references therein).<sup>10</sup>

Dynamic mechanism design is related to the literature on multidimensional screening, as noted, e.g., by Rochet and Stole (2003). Nevertheless, there is a sense in which incentive compatibility is easier to ensure in a dynamic setting than in a static multidimensional setting. This is because in a dynamic setting an agent is asked to report each dimension of his private information before learning the subsequent dimensions, and so has fewer deviations available than in the corresponding static setting in which he observes all the dimensions at once. Because of this, the set of implementable

---

<sup>8</sup>Rahman (2010) derived a general characterization of implementable dynamic allocation rules similar to Rochet's (1987) cyclical monotonicity. Its applicability to the design of optimal mechanisms is, however, yet to be explored.

<sup>9</sup>Building on this literature and on the results of the current paper, Garrett (2011) combines private information about arrival dates with time-varying types.

<sup>10</sup>See Battaglini (2007) and Strulovici (2011) for analysis of limited commitment with changing types.

allocation rules is larger in a dynamic setting than in the corresponding static multidimensional setting. On the other hand, our necessary conditions for incentive compatibility are valid also for multidimensional screening problems.

## 2 The Environment

**Conventions.** For any set  $B$ ,  $B^{-1}$  denotes a singleton. If  $B$  is measurable,  $\Delta(B)$  is the set of probability measures over  $B$ . Any function defined on a measurable set is assumed measurable. Tildes distinguish random variables from realizations so that, for example,  $\theta$  denotes a realization of  $\tilde{\theta}$ . Any set of real vectors or sequences is endowed with the product order unless noted otherwise.

**Decisions.** Time is discrete and indexed by  $t = 0, 1, \dots, \infty$ . There are  $n \geq 1$  agents, indexed by  $i = 1, \dots, n$ . In every period  $t$ , each agent  $i$  observes a signal, or type,  $\theta_{it} \in \Theta_{it} = (\underline{\theta}_{it}, \bar{\theta}_{it}) \subseteq \mathbb{R}$ , with  $-\infty \leq \underline{\theta}_{it} \leq \bar{\theta}_{it} \leq +\infty$ , and then sends a message to a mechanism which leads to an allocation  $x_{it} \in X_{it}$  and a payment  $p_{it} \in \mathbb{R}$  for each agent  $i$ . Each  $X_{it}$  is assumed to be a measurable space (with the sigma-algebra left implicit). The set of feasible allocation sequences is  $X \subseteq \prod_{t=0}^{\infty} \prod_{i=1}^n X_{it}$ . This formulation allows for the possibility that feasible allocations in a given period depend on the allocations in the previous periods, or that the feasible allocations for agent  $i$  depend on the other agents' allocations.<sup>11</sup> Let  $X_t \equiv \prod_{i=1}^n X_{it}$ ,  $X_i^t \equiv \prod_{s=0}^t X_{is}$ , and  $X^t \equiv \prod_{s=0}^t X_s$ . The sets  $\Theta_t$ ,  $\Theta_i^t$ , and  $\Theta^t$  are defined analogously. Let  $\Theta_i^\infty \equiv \prod_{t=0}^{\infty} \Theta_{it}$  and  $\Theta \equiv \prod_{i=1}^n \Theta_i^\infty$ .

In every period  $t$ , each agent  $i$  observes his own allocation  $x_{it}$  but not the other agents' allocations  $x_{-i,t}$ .<sup>12</sup> The observability of  $x_{it}$  should be thought of as a constraint: a mechanism can reveal more information to agent  $i$  than  $x_{it}$ , but cannot conceal  $x_{it}$ . Our necessary conditions for incentive compatibility do not depend on what additional information is disclosed to the agent by the mechanism. Hence it is convenient to assume that the agents do not observe anything beyond  $\theta_{it}$  and  $x_{it}$ , not even their own transfers. (If the horizon is finite, this is without loss as transfers could be postponed until the end.) As for sufficient conditions, we provide conditions under which more information can be disclosed to the agents without violating incentive compatibility. In particular, we construct payments that can be disclosed in each period, and identify conditions under which the other agents' reports and allocations can also be disclosed.

**Types.** The evolution of agent  $i$ 's information is described by a collection of kernels  $F_i \equiv \langle F_{it} : \Theta_i^{t-1} \times X_i^{t-1} \rightarrow \Delta(\Theta_{it}) \rangle_{t=0}^{\infty}$ , where  $F_{it}(\theta_i^{t-1}, x_i^{t-1})$  denotes the distribution of the random vari-

<sup>11</sup>For example, the (intertemporal) allocation of a private good in fixed supply  $\bar{x}$  can be modelled by letting  $X_{it} = \mathbb{R}_+$  and putting  $X = \{x \in \mathbb{R}_+^{\infty N} : \sum_{it} x_{it} \leq \bar{x}\}$ , while the provision of a public good whose period- $t$  production is independent of the level of production in any other period can be modelled by letting  $X = \{x \in \mathbb{R}_+^{\infty N} : x_{1t} = x_{2t} = \dots = x_{Nt} \text{ all } t\}$ .

<sup>12</sup>This formulation does not explicitly allow for decisions that are not observed by any agent at the time they are made; however, such decisions can be accommodated by introducing a fictitious agent observing them.



able  $\tilde{\theta}_{it}$ , given the history of signals  $\theta_i^{t-1} \in \Theta_i^{t-1}$  and allocations  $x_i^{t-1} \in X_i^{t-1}$ . The dependence on past allocations can capture, for example, learning-by-doing or experimentation (see the bandit-auction application in Section 4). The time- $t$  signals of different agents are drawn independently of each other. That is, the vector  $(\tilde{\theta}_{1t}, \dots, \tilde{\theta}_{nt})$  is distributed according to the product measure  $\prod_{i=1}^n F_{it}(\theta_i^{t-1}, x_i^{t-1})$ . We abuse notation by using  $F_{it}(\cdot | \theta_i^{t-1}, x_i^{t-1})$  to denote the cumulative distribution function (c.d.f.) corresponding to the measure  $F_{it}(\theta_i^{t-1}, x_i^{t-1})$ .

Note that we build in the assumption of *independent types* in the sense of Athey and Segal (2013): in addition to independence of agents' signals within any period  $t$ , we require that the distribution of agent  $i$ 's private signal be determined by things he has observed, that is, by  $(\theta_i^{t-1}, x_i^{t-1})$ . Without these restrictions, payoff equivalence in general fails by an argument analogous to that of Cremer and McLean (1988). On the other hand, dependence on other agents' past signals through the implemented observable decisions  $x_i^{t-1}$  is allowed.

**Preferences.** Each agent  $i$  has von Neumann-Morgenstern preferences over lotteries on  $\Theta \times X \times \mathbb{R}^\infty$ , described by a Bernoulli utility function of the quasilinear form  $U_i(\theta, x) + \sum_{t=0}^\infty \delta^t p_{it}$ , where  $U_i : \Theta \times X \rightarrow \mathbb{R}$ , and  $\delta \in (0, 1]$  is a discount factor common to all agents.<sup>13</sup> The special case of “finite horizon” arises when each  $U_i(\theta, x)$  depends only on  $(\theta^T, x^T)$  for some finite  $T$ .

**Choice rules.** A choice rule consists of an allocation rule  $\chi : \Theta \rightarrow X$  and a transfer rule  $\psi : \Theta \rightarrow \mathbb{R}^\infty \times \dots \times \mathbb{R}^\infty$  such that for all  $t \geq 0$ , the allocation  $\chi_t(\theta)$  and transfers  $\psi_t(\theta)$  implemented in period  $t$  depend only on the history  $\theta^t$  (and so will be written as  $\chi_t(\theta^t)$  and  $\psi_t(\theta^t)$ ). We denote the set of feasible allocation rules by  $\mathcal{X}$ . The restriction to deterministic rules is without loss of generality since randomizations can be generated by introducing a fictitious agent and conditioning on his reports. (Below we provide conditions for an optimal allocation rule to be deterministic.)

**Stochastic processes.** Given the kernels  $F \equiv (F_i)_{i=1}^n$ , an allocation rule  $\chi \in \mathcal{X}$  uniquely defines a stochastic process over  $\Theta$ , which we denote by  $\lambda[\chi]$ .<sup>14</sup> For any period  $t \geq 0$  and history  $\theta^t \in \Theta^t$ , we let  $\lambda[\chi] | \theta^t$  denote the analogous process where  $\tilde{\theta}^t$  is first drawn from a degenerate distribution at  $\theta^t$ , and then the continuation process is generated by applying the kernels and the allocation rule starting from the history  $(\theta^t, \chi^t(\theta^t))$ .

When convenient, we view each agent  $i$ 's private information as being generated by his initial signal  $\theta_{i0}$  and a sequence of “independent shocks.” That is, we assume that for each agent  $i$ , there exist a collection  $\langle \mathcal{E}_i, G_i, z_i \rangle$  where  $\mathcal{E}_i \equiv \langle \mathcal{E}_{it} \rangle_{t=0}^\infty$  is a collection of measurable spaces,  $G_i \equiv \langle G_{it} \rangle_{t=0}^\infty$  is a collection of probability distributions with  $G_{it} \in \Delta(\mathcal{E}_{it})$  for  $t \geq 0$ , and  $z_i \equiv \langle z_{it} : \Theta_i^{t-1} \times X_i^{t-1} \times \mathcal{E}_{it} \rightarrow \Theta_{it} \rangle_{t=0}^\infty$  is a sequence of functions such that, for all  $t \geq 0$  and  $(\theta_i^{t-1}, x_i^{t-1}) \in$

<sup>13</sup>As usual, we may alternatively interpret  $p_{it}$  as agent  $i$ 's utility from his period- $t$  payment (See, e.g., Garrett and Pavan (2013) and the discussion below). Furthermore, Theorem 1 below extends as stated to environments where  $i$ 's utility is of the form  $U_i(\theta, x) + P_i(p_{i0}, p_{i1}, \dots)$  for an arbitrary function  $P_i : \mathbb{R}^\infty \rightarrow \mathbb{R}$ . (A model without transfers corresponds to the special case where  $P_i \equiv 0$ , all  $i$ ).

<sup>14</sup>Existence and uniqueness follows by the Tulcea extension theorem (see, e.g., Pollard, 2002, Ch.4, Theorem 49).

$\Theta_i^{t-1} \times X_i^{t-1}$ , if  $\tilde{\varepsilon}_{it}$  is distributed according to  $G_{it}$ , then  $z_{it}(\theta_i^{t-1}, x_i^{t-1}, \tilde{\varepsilon}_{it})$  is distributed according to  $F_{it}(\theta_i^{t-1}, x_i^{t-1})$ . Given any allocation rule  $\chi$ , we can then think of the process  $\lambda[\chi]$  being generated as follows: Let  $\tilde{\varepsilon}$  be distributed on  $\prod_{t=1}^{\infty} \prod_{i=1}^n \mathcal{E}_{it}$  according to the product measure  $\prod_{t=1}^{\infty} \prod_{i=1}^n G_{it}$ . Draw the period-0 signals  $\theta_0$  according to the initial distribution  $\prod_{i=1}^n F_{i0}$  independently of  $\tilde{\varepsilon}$ , and construct types for periods  $t > 0$  recursively by  $\theta_{it} = z_{it}(\theta_i^{t-1}, \chi^{t-1}(\theta^{t-1}), \varepsilon_{it})$ . (Note that we can think of each agent  $i$  observing the shock  $\varepsilon_{it}$  in each period  $t$ , yet  $(\theta_i^t, x_i^{t-1})$  remains a sufficient statistic for his payoff-relevant private information in period  $t$ .) It is a standard result on stochastic processes that such a *state representation*  $\langle \mathcal{E}_i, G_i, z_i \rangle_{i=1}^n$  exists for any kernels  $F$ .<sup>15</sup> For example, if agent  $i$ 's signals follow a linear autoregressive process of order 1, then the  $z_i$  functions take the familiar form  $z_{it}(\theta_i^{t-1}, x_i^{t-1}, \varepsilon_{it}) = \phi_i \theta_{i,t-1} + \varepsilon_{it}$  for some  $\phi_i \in \mathbb{R}$ . The general case can be handled as follows:

**Example 1 (Canonical representation)** Fix the kernels  $F$ . For all  $i = 1, \dots, n$ , and  $t \geq 1$ , let  $\mathcal{E}_{it} = (0, 1)$ , let  $G_{it}$  be the uniform distribution on  $(0, 1)$ , and define the generalized inverse  $F_{it}^{-1}$  by setting  $F_{it}^{-1}(\varepsilon_{it} | \theta_i^{t-1}, x_i^{t-1}) \equiv \inf\{\theta_{it} : F_{it}(\theta_{it} | \theta_i^{t-1}, x_i^{t-1}) \geq \varepsilon_{it}\}$  for all  $\varepsilon_{it} \in (0, 1)$  and  $(\theta_i^{t-1}, x_i^{t-1}) \in \Theta_i^{t-1} \times X_i^{t-1}$ . The random variable  $F_{it}^{-1}(\tilde{\varepsilon}_{it} | \theta_i^{t-1}, x_i^{t-1})$  is then distributed according to the c.d.f.  $F_{it}(\cdot | \theta_i^{t-1}, x_i^{t-1})$  so that we can put  $z_{it}(\theta_i^{t-1}, x_i^{t-1}, \varepsilon_{it}^i) = F_{it}^{-1}(\varepsilon_{it} | \theta_i^{t-1}, x_i^{t-1})$ .<sup>16</sup> We refer to the state representation so defined as the *canonical representation* of  $F$ .  $\square$

Nevertheless, the canonical representation is not always the most convenient as many processes such as the AR(1) above are naturally defined in terms of other representations, and hence we work with the general definition.

In what follows, we use the fact that, given a state representation  $\langle \mathcal{E}_i, G_i, z_i \rangle_{i=1}^n$ , for any period  $s \geq 0$ , each agent  $i$ 's continuation process can be expressed directly in terms of the history  $\theta_i^s$  and shocks  $\varepsilon_{it}$ ,  $t \geq 0$ , by defining the functions  $Z_{i,(s),t} \equiv \langle Z_{i,(s),t} : \Theta_i^s \times X_i^{t-1} \times \mathcal{E}_i^t \rightarrow \Theta_{it} \rangle_{t=0}^{\infty}$  recursively by  $Z_{i,(s),t}(\theta_i^s, x_i^{t-1}, \varepsilon_{it}^t) = z_{it}(Z_{i,(s)}^{t-1}(\theta_i^s, x_i^{t-2}, \varepsilon_{it}^{t-1}), x_i^{t-1}, \varepsilon_{it})$ , where  $Z_{i,(s)}^{t-1}(\theta_i^s, x_i^{t-2}, \varepsilon_{it}^{t-1}) \equiv (Z_{i,(s),\tau}(\theta_i^s, x_i^{\tau-1}, \varepsilon_{it}^{\tau}))_{\tau=0}^{t-1}$  with  $Z_{i,(s),t}(\theta_i^s, x_i^{t-1}, \varepsilon_{it}^t) \equiv \theta_{it}$  for all  $t \leq s$ .

## 2.1 Regularity Conditions

Similarly to static models with continuous types, our analysis requires that each agent's expected utility be sufficiently well-behaved function of his private information. In a dynamic model, an agent's expected continuation utility depends on his current type both directly through the utility function as well as through its impact on the distribution of future types. Hence we impose regularity conditions on both the utility functions and the kernels.

**Condition (U-D) Utility Differentiable:** For all  $i = 1, \dots, n$ ,  $t \geq 0$ ,  $x \in X$ , and  $\theta \in \Theta$ ,  $U_i(\theta_i, \theta_{-i}, x)$  is a differentiable function of  $\theta_i^t \in \Theta_i^t$ .

<sup>15</sup>This observation was first used in a mechanism-design context by Esó and Szentes (2007), who studied a two-period model of information disclosure in auctions.

<sup>16</sup>This construction is standard; see the second proof of Kolmogorov extension theorem in Billingsley (1995, p.490).

With a finite horizon  $T$ , this condition simply means that  $U_i(\theta_i^T, \theta_{-i}^T, x^T)$  is differentiable in  $\theta_i^T$ . Next, define the norm  $\|\cdot\|$  on  $\mathbb{R}^\infty$  by  $\|y\| \equiv \sum_{t=0}^{\infty} \delta^t |y_t|$ , and let  $\Theta_{i\delta} \equiv \{\theta_i \in \Theta_i^\infty : \|\theta_i\| < \infty\}$ .<sup>17</sup>

**Condition (U-ELC) Utility Equi-Lipschitz Continuous:** For all  $i = 1, \dots, n$ , the family  $\{U_i(\cdot, \theta_{-i}, x)\}_{\theta_{-i} \in \Theta_{-i}, x \in X}$  is equi-Lipschitz continuous on  $\Theta_{i\delta}$ . That is, there exists  $A_i \in \mathbb{R}$  such that  $|U_i(\theta'_i, \theta_{-i}, x) - U_i(\theta_i, \theta_{-i}, x)| \leq A_i \|\theta'_i - \theta_i\|$  for all  $\theta_i, \theta'_i \in \Theta_{i\delta}$ ,  $\theta_{-i} \in \Theta_{-i}$ , and  $x \in X$ .

Conditions U-D and U-ELC are roughly analogous to the differentiability and bounded-derivative conditions imposed in static models (c.f., Milgrom and Segal, 2002). For example, stationary payoffs  $U_i(\theta, x) = \sum_{t=0}^{\infty} \delta^t u_i(\theta_t, x_t)$  satisfy U-D and U-ELC if  $u_i$  is differentiable and equi-Lipschitz in  $\theta_{it}$  (e.g., linear payoffs  $u_i(\theta_t, x_t) = \theta_{it} x_{it}$  are fine provided that  $x_{it}$  is bounded).

**Condition (F-BE) Process Bounded in Expectation:** For all  $i = 1, \dots, n$ ,  $t \geq 0$ ,  $\theta^t \in \Theta^t$ , and  $\chi \in \mathcal{X}$ ,  $\mathbb{E}^{\lambda[\chi]|\theta^t} [\|\tilde{\theta}_{i\tau}\|] < \infty$ .

Condition F-BE implies that, for any allocation rule  $\chi$  and any period- $t$  type history  $\theta^t$  ( $t \geq 0$ ), the sequence of agent  $i$ 's future types has a finite norm with  $\lambda[\chi]|\theta^t$ -probability 1. This allows us to effectively restrict attention to the space  $\Theta_{i\delta}$ . With a finite horizon, F-BE simply requires that, for all  $t \geq 0$ , the expectation of each  $\tilde{\theta}_{i\tau}$ , with  $t < \tau \leq T$ , exists conditional on any  $\theta^t$ .

**Condition (F-BIR) Process Bounded Impulse Responses:** There exist a state representation  $\langle \mathcal{E}_i, G_i, z_i \rangle_{i=1}^n$  and functions  $C_{i,(s)} : \mathcal{E}_i \rightarrow \mathbb{R}^\infty$ ,  $i = 1, \dots, n$ ,  $s \geq 0$ , with  $\mathbb{E}[\|C_{i,(s)}(\tilde{\varepsilon}_i)\|] \leq B_i$  for some constant  $B_i$  independent of  $s$ , such that for all  $i = 1, \dots, n$ ,  $t \geq s$ ,  $\theta_i^s \in \Theta_i^s$ ,  $x_i \in X_i$ , and  $\varepsilon_i^t \in \mathcal{E}_i^t$ ,  $Z_{i,(s),t}(\theta_i^s, x_i^{t-1}, \varepsilon_i^t)$  is a differentiable function of  $\theta_{is}$  with  $|\partial Z_{i,(s),t}(\theta_i^s, x_i^{t-1}, \varepsilon_i^t) / \partial \theta_{is}| \leq C_{i,(s),t-s}(\varepsilon_i)$ .

Condition F-BIR is essentially the process-analog of Conditions U-D and U-ELC. It guarantees that small changes in the current type have a small effect on future types. We provide a way to check F-BIR as well as examples of kernels that satisfy it in later sections (see, e.g., Example 3).

Finally, we impose the following bounds on the agents' utility functions to ensure that the expected net present value of the transfers we construct exists when the horizon is infinite:

**Condition (U-SPR) Utility Spreadable:** For all  $i = 1, \dots, n$ , there exists a sequence of functions  $\langle u_{it} : \Theta^t \times X^t \rightarrow \mathbb{R} \rangle_{t=0}^{\infty}$  and constants  $L_i$  and  $(M_{it})_{t=0}^{\infty}$ , with  $L_i, \|M_{it}\| < \infty$ , such that for all  $(\theta, x) \in \Theta \times X$  and  $t \geq 0$ ,  $U_i(\theta, x) = \sum_{t=0}^{\infty} \delta^t u_{it}(\theta^t, x^t)$  and  $|u_{it}(\theta^t, x^t)| \leq L_i |\theta_{it}| + M_{it}$ .

The condition is satisfied, for example, if the functions  $u_{it}$  are uniformly bounded, or take the linear form  $u_{it}(\theta^t, x^t) = \theta_{it} x_{it}$  with  $X_{it}$  bounded (but  $\Theta_{it}$  possibly unbounded).

For ease of reference, we combine the above conditions into a single definition.

---

<sup>17</sup>It is possible to rescale  $\theta_{it}$  and work with the standard  $l_1$  norm. However, we use the weighted norm to deal without rescaling with the standard economic applications with time discounting. Note also that for a finite horizon, the norm  $\|\cdot\|$  is equivalent to the Euclidean norm, and so the choice is irrelevant. For infinite horizon, increasing  $\delta$  weakens the conditions imposed on the utility function while strengthening the conditions imposed on the kernels.

**Definition 1 (Regular environment)** *The environment is **regular** if it satisfies conditions U-D, U-ELC, F-BE, F-BIR, and U-SPR.*

### 3 PBE-Implementability

Following Myerson (1986), we restrict attention to direct mechanisms where, in every period  $t$ , each agent  $i$  confidentially reports a type from his type space  $\Theta_{it}$ , no information is disclosed to him beyond his allocation  $x_{it}$ , and the agents report truthfully on the equilibrium path. Such a mechanism induces a dynamic Bayesian game between the agents, and hence we use perfect Bayesian equilibrium (PBE) as our solution concept.

Formally, a reporting strategy for agent  $i$  is a collection  $\sigma_i \equiv \langle \sigma_{it} : \Theta_i^t \times \Theta_i^{t-1} \times X_i^{t-1} \rightarrow \Theta_{it} \rangle_{t=0}^\infty$ , where  $\sigma_{it}(\theta_i^t, \hat{\theta}_i^{t-1}, x_i^{t-1}) \in \Theta_{it}$  is agent  $i$ 's report in period  $t$  when his true type history is  $\theta_i^t$ , his reported type history is  $\hat{\theta}_i^{t-1}$ , and his allocation history is  $x_i^{t-1}$ . The strategy  $\sigma_i$  is *on-path truthful* if  $\sigma_{it}((\theta_i^{t-1}, \theta_{it}), \hat{\theta}_i^{t-1}, x_i^{t-1}) = \theta_{it}$  for all  $t \geq 1$ ,  $\theta_{it} \in \Theta_{it}$ ,  $\theta_i^{t-1} \in \Theta_i^{t-1}$ , and  $x_i^{t-1} \in X_i^{t-1}$ . Note that there are no restrictions on the behavior that an on-path truthful strategy may prescribe following lies.

The specification of a PBE also includes a belief system  $\Gamma$ , which describes each agent  $i$ 's beliefs at each of his information sets  $(\theta_i^t, \hat{\theta}_i^{t-1}, x_i^{t-1})$  about the unobserved past moves by Nature ( $\theta_{-i}^{t-1}$ ) and by the other agents ( $\hat{\theta}_{-i}^{t-1}$ ). (The agent's beliefs about the contemporaneous types of agents  $j \neq i$  then follow by applying the kernels.) We restrict these beliefs to satisfy two natural conditions:

**B(i)** For all  $i = 1, \dots, n$ ,  $t \geq 0$ , and  $(\theta_i^t, \hat{\theta}_i^{t-1}, x_i^{t-1}) \in \Theta_i^t \times \Theta_i^{t-1} \times X_i^{t-1}$ , agent  $i$ 's beliefs are independent of his true type history  $\theta_i^t$ .

**B(ii)** For all  $i = 1, \dots, n$ ,  $t \geq 0$ , and  $(\theta_i^t, \hat{\theta}_i^{t-1}, x_i^{t-1}) \in \Theta_i^t \times \Theta_i^{t-1} \times X_i^{t-1}$ , agent  $i$ 's beliefs assign probability 1 to the other agents having reported truthfully, i.e., to the event that  $\hat{\theta}_{-i}^{t-1} = \theta_{-i}^{t-1}$ .

Condition B(i) is similar to condition B(i) in Fudenberg and Tirole (1991, p.331). It is motivated by the fact that, given agent  $i$ 's reports  $\hat{\theta}_i^{t-1}$  and observed allocations  $x_i^{t-1}$ , the distribution of his true types  $\theta_i^t$  is independent of the other agents' types or reports. Condition B(ii) in turn says that agent  $i$  always believes that his opponents have been following their equilibrium strategies.<sup>18</sup> Note that under these two conditions, we can describe agent  $i$ 's beliefs as a collection of probability distributions  $\Gamma_{it} : \Theta_{-i}^{t-1} \times X_{-i}^{t-1} \rightarrow \Delta(\Theta_{-i}^{t-1})$ ,  $t \geq 0$ , where  $\Gamma_{it}(\hat{\theta}_{-i}^{t-1}, x_{-i}^{t-1})$  represents agent  $i$ 's beliefs over the other agents' past types  $\theta_{-i}^{t-1}$  (which he believes to be equal to the reports) given that he reported  $\hat{\theta}_i^{t-1}$  and observed the allocations  $x_i^{t-1}$ . We then have the following definitions:

<sup>18</sup>With continuous types, any particular history for agent  $i$  has probability 0, and hence B(ii) cannot be derived from Bayes' rule but has to be imposed. Note that even when the kernels do not have full support, they are defined at all histories, and hence the continuation process is always well defined.

**Definition 2 (On-path truthful PBE; PBIC)** An *on-path truthful PBE* of a direct mechanism  $\langle \chi, \psi \rangle$  is a pair  $(\sigma, \Gamma)$  consisting of an on-path truthful strategy profile  $\sigma$  and a belief system  $\Gamma$  such that (i)  $\Gamma$  satisfies conditions  $B(i)$  and  $B(ii)$  and is consistent with Bayes' rule on all positive-probability events given  $\sigma$ , and (ii) for all  $i = 1, \dots, n$ ,  $\sigma_i$  maximizes agent  $i$ 's expected payoff at each information set given  $\sigma_{-i}$  and  $\Gamma$ .<sup>19</sup> The choice rule  $\langle \chi, \psi \rangle$  is **Perfect Bayesian Incentive Compatible (PBIC)** if the corresponding direct mechanism has an on-path truthful PBE.

Since the strategy profile  $\sigma$  is on-path truthful, part (i) in the above definition depends only on the allocation rule  $\chi$ , and hence we write  $\mathbf{\Gamma}(\chi)$  for the set of belief systems satisfying part (i). (Each element of  $\mathbf{\Gamma}(\chi)$  is a system of regular conditional probability distributions, the existence of which is well known; see, for example, Dudley, 2002.) Note that the concept of PBIC implies, in particular, that truthful reporting is optimal at every truthful history.

### 3.1 First-Order Necessary Conditions

We start by deriving a necessary condition for PBIC by applying an envelope theorem to an agent's problem of choosing an optimal reporting strategy at an arbitrary truthful history.

Fix a choice rule  $\langle \chi, \psi \rangle$  and a belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$ . Suppose agent  $i$  plays according to an on-path truthful strategy, and consider a period- $t$  history of the form  $((\theta_i^{t-1}, \theta_{it}), \theta_i^{t-1}, \chi_i^{t-1}(\theta^{t-1}))$ , that is, when agent  $i$  has reported truthfully in the past, the complete reporting history is  $\theta^{t-1}$ , and agent  $i$ 's current type is  $\theta_{it}$ . Agent  $i$ 's expected payoff is then given by

$$V_{it}^{\langle \chi, \psi \rangle, \Gamma}(\theta^{t-1}, \theta_{it}) \equiv \mathbb{E}^{\lambda_i[\chi, \Gamma]|\theta^{t-1}, \theta_{it}} \left[ U_i(\chi(\tilde{\theta}), \tilde{\theta}) + \sum_{t=0}^{\infty} \delta^t \psi_{it}(\tilde{\theta}) \right],$$

where  $\lambda_i[\chi, \Gamma]|\theta^{t-1}, \theta_{it}$  is the stochastic process over  $\Theta$  from the perspective of agent  $i$ . Formally,  $\lambda_i[\chi, \Gamma]|\theta^{t-1}, \theta_{it}$  is the unique probability measure on  $\Theta$  obtained by first drawing  $\theta_{-i}^{t-1}$  according to agent  $i$ 's belief  $\Gamma_{it}(\theta_i^{t-1}, \chi_i^{t-1}(\theta^{t-1}))$ , drawing  $\theta_{-i,t}$  according to  $\prod_{j \neq i} F_{jt}(\theta_j^{t-1}, \chi_j^{t-1}(\theta_i^{t-1}, \theta_{-i}^{t-1}))$ , and then using the allocation rule  $\chi$  and the kernels  $F$  to generate the process from period- $t$  onwards. Note that in period 0, this measure is only a function of the kernels, and hence we write it as  $\lambda_i[\chi]|\theta_{i0}$ , and similarly omit the belief system  $\Gamma$  in  $V_{i0}^{\langle \chi, \psi \rangle}(\theta_{i0})$ .

The following is a dynamic version of the envelope condition familiar from static models:

**Definition 3 (ICFOC)** Fix  $i = 1, \dots, n$  and  $s \geq 0$ . The choice rule  $\langle \chi, \psi \rangle$  with belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$  satisfies **ICFOC** $_{i,s}$  if, for all  $\theta^{s-1} \in \Theta^{s-1}$ ,  $V_{is}^{\langle \chi, \psi \rangle, \Gamma}(\theta^{s-1}, \theta_{is})$  is a Lipschitz continuous function of  $\theta_{is}$  with the derivative given a.e. by

$$\frac{\partial V_{is}^{\langle \chi, \psi \rangle, \Gamma}(\theta^{s-1}, \theta_{is})}{\partial \theta_{is}} = \mathbb{E}^{\lambda_i[\chi, \Gamma]|\theta^{s-1}, \theta_{is}} \left[ \sum_{t=s}^{\infty} I_{i,(s),t}(\tilde{\theta}_i^t, \chi_i^{t-1}(\tilde{\theta})) \frac{\partial U_i(\tilde{\theta}, \chi(\tilde{\theta}))}{\partial \theta_{it}} \right], \quad (1)$$

<sup>19</sup>In particular, the expected allocation utility and the expected net present value of transfers from an on-path truthful strategy are well-defined and finite conditional on any truthful history.

where<sup>20</sup>

$$I_{i,(s),t}(\theta_i^t, x_i^{t-1}) \equiv \mathbb{E} \left[ \frac{\partial Z_{i,(s),t}(\theta_i^s, x_i^{t-1}, \tilde{\varepsilon}_i^t)}{\partial \theta_{is}} \middle| Z_{i,(s),t}^t(\theta_i^s, x_i^{t-1}, \tilde{\varepsilon}_i^t) = \theta_i^t \right]. \quad (2)$$

The choice rule  $\langle \chi, \psi \rangle$  satisfies **ICFOC** if there exists a belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$  such that  $\langle \chi, \psi \rangle$  with belief system  $\Gamma$  satisfies  $\text{ICFOC}_{i,s}$  for all agents  $i = 1, \dots, n$  and all periods  $s \geq 0$ .

**Theorem 1** Suppose the environment is regular.<sup>21</sup> Then every PBIC choice rule satisfies ICFOC.

By Theorem 1, the formula in (1) is a dynamic generalization of the envelope formula familiar from static mechanism design (Mirrlees, 1971, Myerson, 1981). By inspection, the period-0 formula implies a weak form of *dynamic payoff (and revenue) equivalence*: each agent's period-0 interim expected payoff is pinned down by the allocation rule  $\chi$  up to a constant. (We provide a stronger payoff-equivalence result in the next section.) Special cases of the envelope formula (1) have been identified by Baron and Besanko (1984), Besanko (1985), Courty and Li (2000), and Esó and Szentes (2007), among others. However, it should be noted that the contribution of Theorem 1 is not just in generalizing the formula, but in providing conditions on the utility functions and type processes that imply that ICFOC is indeed a necessary condition for all PBIC choice rules.

Heuristically, the proof of Theorem 1 in the Appendix proceeds by applying an envelope-theorem argument to the normal form of an agent's problem of choosing an optimal continuation strategy at a given truthful history. The argument is identical across agents and periods, and hence without loss of generality we focus on establishing  $\text{ICFOC}_{i,0}$  for some agent  $i$  by considering his period-0 ex-interim problem of choosing a reporting strategy conditional on his initial signal  $\theta_{i0}$ . Nevertheless, Theorem 1 is not an immediate corollary of Milgrom and Segal's (2002) envelope theorem for arbitrary choice sets. Namely, their result requires that the objective function be differentiable in the parameter (with an appropriately bounded derivative) for *any* feasible element of the choice set. Here it would require that, for any initial report  $\hat{\theta}_{i0}$  and any plan  $\rho \equiv \langle \rho_t : \prod_{\tau=1}^t \Theta_\tau \rightarrow \Theta_t \rangle_{t=1}^\infty$  for reporting future signals, agent  $i$ 's payoff be differentiable in  $\theta_{i0}$ . But this property need not hold in a regular environment. This is because a change in the initial signal  $\theta_{i0}$  changes the distribution of the agent's future signals, which in turn changes the distribution of his future reports and allocations through the plan  $\rho$  and the choice rule  $\langle \chi, \psi \rangle$ . For some combinations of  $\hat{\theta}_0$ ,  $\rho$ , and  $\langle \chi, \psi \rangle$ , this may lead to an expected payoff that is non-differentiable or even discontinuous in  $\theta_{i0}$ .<sup>22</sup>

<sup>20</sup>The  $I_{i,(s),t}$  functions are conditional expectations and thus defined only up to sets of measure 0.

<sup>21</sup>Condition U-SPR, which requires utility to be spreadable, is not used in the proof of this theorem.

<sup>22</sup>For a simple example, consider a two-period environment with one agent and a single indivisible good to be allocated in the second period as in Courty and Li (2000) (i.e.,  $X_0 = \{0\}$ ,  $X_1 = \{0, 1\}$ ). Suppose the agent's payoff is of the form  $U(\theta, x) = \theta_1 x_1$ , and that  $\tilde{\theta}_0$  is distributed uniformly on  $(0, 1)$  with  $\tilde{\theta}_1 = \tilde{\theta}_0$  almost surely. (It is straightforward to verify that this environment is regular; e.g., put  $Z_1(\theta_0, \varepsilon_1) = \theta_0$  for all  $(\theta_0, \varepsilon_1)$  to verify F-BIR). Consider the PBIC choice rule  $\langle \chi, \psi \rangle$  where  $\chi$  is defined by  $\chi_0 = 0$ ,  $\chi_1(\theta_0, \theta_1) = 1$  if  $\theta_0 = \theta_1 \geq \frac{1}{2}$ , and  $\chi_1(\theta_0, \theta_1) = 0$  otherwise, and where  $\psi$  is defined by setting  $\psi_0 = 0$  and  $\psi_1(\theta_0, \theta_1) = \frac{1}{2} \chi_1(\theta_0, \theta_1)$ . Now, fix an initial report  $\hat{\theta}_0 > \frac{1}{2}$  and fix the plan  $\rho$  that

To deal with this complication, we transform the problem into one where the distribution of agent  $i$ 's future information is independent of his initial signal  $\theta_{i0}$  so that changing  $\theta_{i0}$  leaves future reports unaltered.<sup>23</sup> This is done by using a state representation to generate the signal process and by asking the agent to report his initial signal  $\theta_{i0}$  and his future *shocks*  $\varepsilon_{it}$ ,  $t \geq 1$ . This equivalent formulation provides an additional simplification in that we may assume that agent  $i$  reports the shocks truthfully.<sup>24</sup> The rest of the proof then amounts to showing that if the environment is regular, then the transformed problem is sufficiently well-behaved to apply arguments similar to those in Milgrom and Segal (2002).

**Remark 1** If the notion of incentive compatibility is weakened from PBIC to the requirement that on-path truthful strategies are a Bayesian-Nash equilibrium of the direct mechanism, then the above argument can still be applied in period 0 to establish  $\text{ICFOC}_{i,0}$  for all  $i = 1, \dots, n$ . Thus the weak payoff equivalence discussed above holds across all Bayesian-Nash incentive compatible mechanisms.  $\square$

We finish this subsection with two examples that suggest an interpretation of the functions defined by (2) and establish a connection to the literature. To simplify notation, we restrict attention to the case of a single agent and omit the subscript  $i$ .

**Example 2 (AR(k) process)** Consider the case of a single agent, and suppose that the signal  $\theta_t$  evolves according to an autoregressive (AR) process that is independent of the allocations:

$$\tilde{\theta}_t = \sum_{j=1}^{\infty} \phi_j \tilde{\theta}_{t-j} + \tilde{\varepsilon}_t,$$

where  $\tilde{\theta}_t = 0$  for all  $t < 0$ ,  $\phi_j \in \mathbb{R}$  for all  $j \in \mathbb{N}$ , and  $\tilde{\varepsilon}_t$  is a random variable distributed according to some c.d.f.  $G_t$  with support  $\mathcal{E}_t \subseteq \mathbb{R}$ , with all the  $\tilde{\varepsilon}_t$ ,  $t \geq 0$ , distributed independently of each other and of  $\tilde{\theta}_0$ . Note that we have defined the process in terms of the state representation  $\langle \mathcal{E}, G, z \rangle$ , where  $z_t(\theta^{t-1}, x^{t-1}, \varepsilon_t) = \sum_{j=1}^{\infty} \phi_j \theta_{t-j} + \varepsilon_t$ . The functions (2) are then time-varying scalars

$$I_{(s),s} = 1 \text{ and } I_{(s),t} = \frac{\partial Z_{(s),t}}{\partial \theta_s} = \sum_{\substack{K \in \mathbb{N}, l \in \mathbb{N}^{K+1}: \\ s=l_0 < \dots < l_K=t}} \prod_{k=1}^K \phi_{l_k - l_{k-1}} \text{ for } t > s. \quad (3)$$

---

reports  $\theta_1$  truthfully in period 1 (i.e.,  $\rho(\theta_1) = \theta_1$  for all  $\theta_1$ ). The resulting expected payoff is  $\hat{\theta}_0 - \frac{1}{2} > 0$  for  $\theta_0 = \hat{\theta}_0$ , whereas it is equal to 0 for all  $\theta_0 \neq \hat{\theta}_0$ . That is, the expected payoff is discontinuous in the true initial type at  $\theta_0 = \hat{\theta}_0$ .

<sup>23</sup>In an earlier draft we showed that, when the horizon is finite, the complication can alternatively be dealt with by using backward induction. Roughly, this solves the problem as it forces the agent to use a sequentially rational continuation strategy given any initial report, and thus rules out problematic elements of his feasible set.

<sup>24</sup>By PBIC, truthful reporting remains optimal in the restricted problem where the agent can only choose  $\hat{\theta}_0$ , and hence the value function that we are trying to characterize is unaffected. (In terms of the kernel representation, this amounts to restricting each type  $\theta_{i0}$  to using strategies where given any initial report  $\hat{\theta}_{i0} \in \Theta_{i0}$ , the agent is constrained to report  $\hat{\theta}_{it} = Z_{it}(\hat{\theta}_{i0}, x_i^{t-1}, \varepsilon_i^t)$  in period  $t$ .) Note that restricting the agent to report truthfully his future  $\theta_{it}$  would not work as the resulting restricted problem is not sufficiently well-behaved in general; see footnote 22.

In the special case of an AR(1) process, as in Besanko (1985), we have  $\phi_j = 0$  for all  $j > 1$ , and hence the formula simplifies to  $I_{(s),t} = (\phi_1)^{t-s}$ . Condition F-BIR requires that there exist  $B \in \mathbb{R}$  such that  $\|I_{(s)}\| \equiv \sum_{t=0}^{\infty} \delta^t |I_{(s),t}| < B$  for all  $s \geq 0$ , which in the AR(1) case is satisfied if and only if  $\delta |\phi_1| < 1$ . For Condition F-BE, write

$$\begin{aligned} \tilde{\theta}_t &= Z_{(0),t}(\theta_0, \tilde{\varepsilon}^t) = I_{(0),t}\theta_0 + \sum_{\tau=1}^t I_{(0),t-\tau}\tilde{\varepsilon}_\tau \text{ for all } t \geq 0, \text{ so that} \\ \mathbb{E}^{\lambda|\theta_0} \left[ \|\tilde{\theta}\| \right] &\leq \|I_{(0)}\| |\theta_0| + \sum_{t=1}^{\infty} \delta^t \sum_{\tau=1}^t |I_{(0),t-\tau}| \mathbb{E} \left[ \|\tilde{\varepsilon}_\tau\| \right] = \|I_{(0)}\| (|\theta_0| + \mathbb{E} \left[ \|\tilde{\varepsilon}\| \right]). \end{aligned}$$

Similarly, we have  $\mathbb{E}^{\lambda|\theta^s} \left[ \|\tilde{\theta}\| \right] \leq \sum_{m=0}^{s-1} \|I_{(m)}\| |\theta_m| + \|I_{(s)}\| (|\theta_s| + \delta^{-s} \mathbb{E} \left[ \|\tilde{\varepsilon}\| \right])$ . Hence, F-BE is ensured by assuming, in addition to the bound  $B$  needed for F-BIR, that  $\mathbb{E} \left[ \|\tilde{\varepsilon}\| \right] < \infty$ , which simply requires that the mean of the shocks grows slower than the discount rate. (E.g., it is trivially satisfied if  $\varepsilon_t$  are i.i.d. with a finite mean.)  $\square$

The constants defined by (3) coincide with the impulse responses of a linear AR process. More generally, the  $I_{(s),t}$  functions in (2) can be interpreted as *nonlinear impulse responses*. To see this, apply Theorem 1 to a regular single-agent environment with fixed decisions and no payments (i.e., with  $X_t = \{\hat{x}_t\}$  and  $\psi_t(\theta) = 0$  for all  $t \geq 0$  and  $\theta \in \Theta$ ), in which case optimization is irrelevant, and we simply have  $V_s^{(\chi, \Psi)}(\theta^s) \equiv \mathbb{E}^{\lambda|\theta^s} [U(\tilde{\theta}, \hat{x})]$ . Then (1) takes the form

$$\frac{d\mathbb{E}^{\lambda|\theta^s} [U(\tilde{\theta}, \hat{x})]}{d\theta_s} = \mathbb{E}^{\lambda|\theta^s} \left[ \sum_{t=s}^{\infty} I_{(s),t}(\tilde{\theta}^t, \hat{x}) \frac{\partial U(\tilde{\theta}, \hat{x})}{\partial \theta_t} \right].$$

Note that the impulse response functions  $I_{(s),t}$  are determined entirely by the stochastic process and satisfy the above equation for any utility function  $U$  satisfying Conditions U-D and U-ELC.<sup>25</sup>

If for all  $t \geq 1$ , the function  $z_t$  in the state representation of the type process is differentiable in  $\theta^{t-1}$ , we can use the chain rule to inductively calculate the impulse responses as

$$\frac{\partial Z_{(s),t}(\theta^s, x^{t-1}, \varepsilon^t)}{\partial \theta_s} = \sum_{\substack{K \in \mathbb{N}, \\ s=l_0 < \dots < l_K=t}} \prod_{l \in \mathbb{N}^{K+1}, k=1}^K \frac{\partial z_{l_k}(Z^{l_k-1}(\theta^s, x^{l_k-2}, \varepsilon^{l_k-1}), x^{l_k-1}, \varepsilon_{l_k})}{\partial \theta_{l_{k-1}}}. \quad (4)$$

The derivative  $\partial z_m / \partial \theta_l$  can be interpreted as the “direct impulse response” of the signal in period  $m$  to the signal in period  $l < m$ . The “total” impulse response  $\partial Z_{(s),t} / \partial \theta_s$  is then obtained by adding up the products of the direct impulse responses over all possible causation chains from period  $s$  to period  $t$ . Applying this observation to the canonical representation yields a simple formula for the impulse responses and a possible way to verify that the kernels satisfy Condition F-BIR:

**Example 3 (Canonical impulse responses)** Suppose that, for all  $t \geq 1$  and  $x^{t-1} \in X^{t-1}$ , the c.d.f.  $F_t(\theta_t | \theta^{t-1}, x^{t-1})$  is continuously differentiable in  $(\theta_t, \theta^{t-1})$ , and let  $f_t(\cdot | \theta^{t-1}, x^{t-1})$  denote the

<sup>25</sup>We conjecture that this property uniquely defines the impulse response functions with  $\lambda|\theta^s$ -probability 1.



density of  $F_t(\cdot|\theta^{t-1}, x^{t-1})$ . Then the direct impulse responses in the canonical representation of Example 1 take the form, for  $(\theta^{m-1}, x^{m-1}, \varepsilon_m) \in \Theta^{m-1} \times X^{m-1} \times (0, 1)$ , and  $m \geq l \geq 0$ ,

$$\frac{\partial z_m(\theta^{m-1}, x^{m-1}, \varepsilon_m)}{\partial \theta_l} = - \left. \frac{\partial F_m(\theta_m|\theta^{m-1}, x^{m-1})/\partial \theta_l}{f_m(\theta_m|\theta^{m-1}, x^{m-1})} \right|_{\theta_m = F_m^{-1}(\varepsilon_m|\theta^{m-1}, x^{m-1})},$$

where we have used the implicit function theorem. Plugging this into equation (4) yields a formula for the impulse responses directly in terms of the kernels. For example, if the agent's type evolves according to a Markov process whose kernels are independent of decisions, the formula simplifies to

$$I_{(s),t}(\theta^t) = \prod_{\tau=s}^t \left( - \frac{\partial F_\tau(\theta_\tau|\theta_{\tau-1})/\partial \theta_{\tau-1}}{f_\tau(\theta_\tau|\theta_{\tau-1})} \right), \quad (5)$$

because then the only causation chain passes through all periods. Two-period versions of this formula appear in Baron and Besanko (1984), Courty and Li (2000), and Esó and Szentes (2007).

As for Condition F-BIR, because the canonical impulse responses are directly in terms of the kernels  $F$ , it is straightforward to back out conditions that guarantee the existence of the bounding functions  $C_{(s)} : \mathcal{E} \rightarrow \mathbb{R}^\infty$ ,  $s \geq 0$ . For example, in the case of a Markov process, it is grossly sufficient that there exists a sequence  $y \in \Theta_\delta$  such that for all  $t \geq 0$ ,  $(\theta_{t-1}, \theta_t) \in \Theta_{t-1} \times \Theta_t$ , and  $x^{t-1} \in X^{t-1}$ , we have  $\left| \frac{\partial F_t(\theta_t|\theta_{t-1}, x^{t-1})/\partial \theta_{t-1}}{f_t(\theta_t|\theta_{t-1}, x^{t-1})} \right| \leq y_t$ . The general case can be handled similarly.  $\square$

**Remark 2** Baron and Besanko (1984) suggested interpreting  $I_{(0),1}(\theta_0, \theta_1) = - \frac{\partial F_1(\theta_1|\theta_0)/\partial \theta_0}{f_1(\theta_1|\theta_0)}$  as a measure of ‘‘informativeness’’ of  $\theta_0$  about  $\theta_1$ . We find the term ‘‘impulse response’’ preferable. First, for linear processes, it matches the usage in the time-series literature. Second, it is more precise. For example, in the two-period case, if  $\tilde{\theta}_1 = \tilde{\theta}_0 + \tilde{\varepsilon}_1$  with  $\tilde{\varepsilon}_1$  normally distributed with mean zero, then the impulse response is identical to 1 regardless of the variance of  $\tilde{\varepsilon}_1$ . On the other hand,  $\theta_0$  is more informative about  $\theta_1$  (in the sense of Blackwell) the smaller the variance of  $\tilde{\varepsilon}_1$ .  $\square$

### 3.2 Payment Construction and Equivalence

Similarly to static settings, for any allocation rule (and belief system), it is possible to use the envelope formula to construct transfers that satisfy first-order conditions at all truthful histories: Fix an allocation rule  $\chi$  and a belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$ . For all  $i = 1, \dots, n$ ,  $s \geq 0$ , and  $\theta \in \Theta$ , let

$$D_{is}^{\chi, \Gamma}(\theta^{s-1}, \theta_{is}) \equiv \mathbb{E}^{\lambda_i[\chi, \Gamma]|\theta^{s-1}, \theta_{is}} \left[ \sum_{t=s}^{\infty} I_{i,(s),t}(\tilde{\theta}_i^t, \chi_i^{t-1}(\tilde{\theta})) \frac{\partial U_i(\tilde{\theta}, \chi(\tilde{\theta}))}{\partial \theta_{it}} \right], \quad (6)$$

$$\text{and } Q_{is}^{\chi, \Gamma}(\theta^{s-1}, \theta_{is}) \equiv \int_{\theta'_{is}}^{\theta_{is}} D_{is}^{\chi, \Gamma}(\theta^{s-1}, q) dq,$$

where  $\theta'_i \in \Theta_{i\delta}$  is some arbitrary fixed type sequence. Define the transfer rule  $\psi$  by setting, for all  $i = 1, \dots, n$ ,  $t \geq 0$ , and  $\theta^t \in \Theta^t$ ,

$$\psi_{it}(\theta^t) = \delta^{-t} Q_{it}^{\chi, \Gamma}(\theta^{t-1}, \theta_{it}) - \delta^{-t} \mathbb{E}^{\lambda_i[\chi, \Gamma]|\theta^{t-1}, \theta_{it}} \left[ Q_{i,t+1}^{\chi, \Gamma}(\tilde{\theta}^t, \tilde{\theta}_{i,t+1}) \right] - \mathbb{E}^{\lambda_i[\chi, \Gamma]|\theta^{t-1}, \theta_{it}} \left[ u_{it}(\tilde{\theta}^t, \chi^t(\tilde{\theta}^t)) \right]. \quad (7)$$

Recall that by Theorem 1, if  $\langle \chi, \psi \rangle$  is PBIC, then agent  $i$ 's expected equilibrium payoff in period  $s$  satisfies  $\partial V_{is}^{\langle \chi, \psi \rangle, \Gamma}(\theta^{s-1}, \theta_{is}) / \partial \theta_{is} = D_{is}^{\chi, \Gamma}(\theta^{s-1}, \theta_{is})$ . Hence the transfer in (7) can be interpreted as agent  $i$ 's information rent (over type  $\theta'_i$ ) as perceived in period  $t$ , net of the rent he expects from the next period, and net of the expected flow utility. We show below that these transfers satisfy ICFOC. In order to address their uniqueness, we introduce the following condition.

**Definition 4 (No leakage)** *The allocation rule  $\chi$  **leaks no information to agent  $i$**  if for all  $t \geq 0$ , and  $\theta_i^{t-1}, \hat{\theta}_i^{t-1} \in \Theta_i^{t-1}$ , the distribution  $F_{it}(\theta_i^{t-1}, \chi_i^{t-1}(\hat{\theta}_i^{t-1}, \theta_{-i}^{t-1}))$  does not depend on  $\theta_{-i}^{t-1}$  (and hence can be written as  $\hat{F}_{it}(\theta_i^{t-1}, \hat{\theta}_i^{t-1})$ ).*

This condition means that observing  $\theta_{it}$  never gives agent  $i$  any information about the other agents' types. Clearly, all allocation rules satisfy it when agent  $i$ 's type evolves autonomously from allocations, or (trivially) in a single-agent setting. We then have our second main result:<sup>26</sup>

**Theorem 2** *Suppose the environment is regular. Then the following statements are true:*

(i) *Given an allocation rule  $\chi$  and a belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$ , let  $\psi$  be the transfer rule defined by (7). Then the choice rule  $\langle \chi, \psi \rangle$  satisfies ICFOC, and for all  $i = 1, \dots, n$ ,  $s \geq 0$ ,  $\theta^{s-1} \in \Theta^{s-1}$ , and  $\theta_{is} \in \Theta_{is}$ ,  $\mathbb{E}^{\lambda_i[\chi, \Gamma]|\theta^{s-1}, \theta_{is}} [|\psi_i(\tilde{\theta})|] < \infty$ .*

(ii) *Let  $\chi$  be an allocation rule that leaks no information to agent  $i$ , and let  $\psi$  and  $\bar{\psi}$  be transfer rules such that the choice rules  $\langle \chi, \psi \rangle$  and  $\langle \chi, \bar{\psi} \rangle$  are PBIC. Then there exists a constant  $K_i$  such that for  $\lambda[\chi]$ -almost every  $\theta_i$ ,*

$$\mathbb{E}^{\lambda_i[\chi]|\theta_i} \left[ \sum_{t=0}^{\infty} \delta^t \psi_{it}(\theta_i, \tilde{\theta}_{-i}) \right] = \mathbb{E}^{\lambda_i[\chi]|\theta_i} \left[ \sum_{t=0}^{\infty} \delta^t \bar{\psi}_{it}(\theta_i, \tilde{\theta}_{-i}) \right] + K_i.$$

**Remark 3** The flow payments  $\psi_{it}(\theta^t)$  defined by (7) are measurable with respect to  $(\theta_i^t, \chi_i^t(\theta^t))$ . Thus, they do not reveal to agent  $i$  any information beyond that contained in the allocations  $x_i$ . Hence they can be disclosed to the agent without affecting his beliefs or incentives.  $\square$

As noted after Theorem 1, ICFOC $_{i,0}$  immediately pins down, up to a constant, the expected net present value of payments  $\mathbb{E}^{\lambda_i[\chi]|\theta_{i0}} \left[ \sum_{t=0}^{\infty} \delta^t \psi_{it}(\tilde{\theta}) \right]$  for each initial type  $\theta_{i0}$  of each agent  $i$  in any PBIC mechanism implementing the allocation rule  $\chi$ . This extends the celebrated revenue equivalence theorem of Myerson (1981) to dynamic environments. Part (ii) of Theorem 2 strengthens the result further to a form of ex-post equivalence. The result is particularly sharp when there is just one agent. Then the no-leakage condition is vacuously satisfied, and the net present value of transfers that implement a given allocation rule  $\chi$  is pinned down up to a single constant *with probability 1* (that

<sup>26</sup>The notation  $\lambda_i[\chi]|\theta_i$  in the proposition denotes the unique measure over the other agents' types  $\Theta_{-i}$  that is obtained from the kernels  $F$  and the allocation  $\chi$  by fixing agent  $i$ 's reports at  $\hat{\theta}_i = \theta_i$ .

is, if  $\langle \chi, \psi \rangle$  and  $\langle \chi, \bar{\psi} \rangle$  are PBIC, then there exists  $K \in \mathbb{R}$  such that  $\sum_{t=0}^{\infty} \delta^t \psi_t(\theta) = \sum_{t=0}^{\infty} \delta^t \bar{\psi}_t(\theta) + K$  for  $\lambda[\chi]$ -almost every  $\theta$ .

The ex-post equivalence of payments is useful for solving mechanism design problems in which the principal cares not just about the expected net present value of payments, but also about how the payments vary with the state  $\theta$  or over time. For example, this includes settings where  $\psi_t(\theta)$  is interpreted as the “utility payment” to the agent in period  $t$ , whose monetary cost to the principal is  $\gamma(\psi_t(\theta))$  for some function  $\gamma$ , as in models with a risk-averse agent. In such models, knowing the net present value of the “utility payments” required to implement a given allocation rule allows computing the cost-minimizing distribution of monetary payments over time (see, for example, Farhi and Werning (2013), or Garrett and Pavan (2013)).

### 3.3 A Characterization for Markov Environments

In order to provide necessary and sufficient conditions for PBE-implementability, we focus on Markov environments, defined formally as follows:

**Definition 5 (Markov environment)** *The environment is **Markov** if, for all  $i = 1, \dots, n$ , the following conditions hold:*

- (i) *Agent  $i$ 's utility function  $U_i$  takes the form  $U_i(\theta, x) = \sum_{t=0}^{\infty} \delta^t u_{it}(\theta_t, x^t)$ .*
- (ii) *For all  $t \geq 1$  and  $x_i^{t-1} \in X_i^{t-1}$ , the distribution  $F_{it}(\theta_i^{t-1}, x_i^{t-1})$  depends on  $\theta_i^{t-1}$  only through  $\theta_{i,t-1}$  (and is hence denoted by  $F_{it}(\theta_{i,t-1}, x_i^{t-1})$ ), and there exist constants  $\phi_i$  and  $(E_{it})_{t=0}^{\infty}$ , with  $\delta\phi_i < 1$  and  $\|E_i\| < \infty$ , such that for all  $(\theta_{it}, x_i^t) \in \Theta_{it} \times X_i^t$ ,  $\mathbb{E}^{F_{i,t+1}(\theta_{it}, x_i^t)} [|\tilde{\theta}_{t+1}|] \leq \phi_i |\theta_{it}| + E_{i,t+1}$ .*

This definition implies that each agent  $i$ 's type process is a Markov decision process, and that his vNM preferences over future lotteries depend on his type history  $\theta_i^t$  only through  $\theta_{it}$  (but can depend on past decisions  $x^{t-1}$ ). The strengthening of Condition F-BE embedded in part (ii) of the definition allows us to establish an appropriate version of the one-stage-deviation principle for the model. Note that every Markov process satisfies the bounds if the sets  $\Theta_{it}$  are bounded.

The key simplification afforded by the Markov assumption is that in a Markov environment, an agent's reporting incentives in any period  $t$  depend only on his current true type and his past reports, but not on his past true types. In particular, if it is optimal for the agent to report truthfully when past reports have been truthful (as in an on-path truthful PBE), then it is also optimal for him to report truthfully even if he has lied in the past. This implies that we can restrict attention to PBE in *strongly truthful strategies*, that is, in strategies that report truthfully at all histories.

We say that the allocation rule  $\chi \in \mathcal{X}$  is *PBE-implementable* if there exists a transfer rule  $\psi$  such that the direct mechanism  $\langle \chi, \psi \rangle$  has an on-path truthful PBE (i.e., it is PBIC). Respectively, we say that  $\chi$  is *strongly PBE-implementable* if there exists a transfer rule  $\psi$  such that the direct mechanism  $\langle \chi, \psi \rangle$  has a strongly truthful PBE. Given an allocation rule  $\chi$ , for all  $i = 1, \dots, n$ ,  $t \geq 0$ ,

and  $\hat{\theta}_{it} \in \Theta_{it}$ , we let  $\chi \circ \hat{\theta}_{it}$  denote the allocation rule obtained from  $\chi$  by replacing  $\theta_{it}$  with  $\hat{\theta}_{it}$  (that is,  $(\chi \circ \hat{\theta}_{it})(\theta) = \chi(\hat{\theta}_{it}, \theta_{i,-t}, \theta_{-i})$  for all  $\theta \in \Theta$ ). Finally, we recall the definition of the  $D$  functions in (6), and recall from Theorem 1 that  $D_{it}^{\chi, \Gamma}(\theta^{t-1}, \theta_{it})$  equals the derivative of agent  $i$ 's expected equilibrium payoff with respect to his current type at any truthful period- $t$  history  $(\theta_i^t, \theta_i^{t-1}, \chi_i^{t-1}(\theta^{t-1}))$  in a PBIC choice rule with allocation rule  $\chi$  and belief system  $\Gamma$ . We then have our third main result:

**Theorem 3** *Suppose the environment is regular and Markov. An allocation rule  $\chi \in \mathcal{X}$  is PBE-implementable if and only if, there exists a belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$  such that for all  $i = 1, \dots, n$ ,  $t \geq 0$ ,  $\theta_{it}, \hat{\theta}_{it} \in \Theta_{it}$ , and  $\theta^{t-1} \in \Theta^{t-1}$ , the following integral monotonicity condition holds:*

$$\int_{\hat{\theta}_{it}}^{\theta_{it}} \left[ D_{it}^{\chi, \Gamma}(\theta^{t-1}, r) - D_{it}^{\chi \circ \hat{\theta}_{it}, \Gamma}(\theta^{t-1}, r) \right] dr \geq 0. \quad (8)$$

When this is the case,  $\chi$  is strongly PBE-implementable with payments given by (7).

The static version of Theorem 3 has appeared in the literature on implementability (see Rochet (1987) or Carbajal and Ely (2013) and the references therein). The basic idea in our proof of the dynamic version is to show that, in every period  $t$ , each agent  $i$ 's problem of reporting a current type is sufficiently well-behaved to allow applying the following simple lemma:

**Lemma 1** *Consider a function  $\Phi : (\underline{\theta}, \bar{\theta})^2 \rightarrow \mathbb{R}$ . Suppose that (a) for all  $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$ ,  $\Phi(\theta, \hat{\theta})$  is a Lipschitz continuous function of  $\theta$ , and (b)  $\bar{\Phi}(\theta) \equiv \Phi(\theta, \theta)$  is a Lipschitz continuous function of  $\theta$ . Then  $\bar{\Phi}(\theta) \geq \Phi(\theta, \hat{\theta})$  for all  $(\theta, \hat{\theta}) \in (\underline{\theta}, \bar{\theta})^2$  if and only if, for all  $(\theta, \hat{\theta}) \in (\underline{\theta}, \bar{\theta})^2$ ,*

$$\int_{\hat{\theta}}^{\theta} \left[ \bar{\Phi}'(q) - \frac{\partial \Phi(q, \hat{\theta})}{\partial \theta} \right] dq \geq 0. \quad (9)$$

**Proof.** For all  $\theta, \hat{\theta} \in (\underline{\theta}, \bar{\theta})$ , let  $g(\theta, \hat{\theta}) \equiv \bar{\Phi}(\theta) - \Phi(\theta, \hat{\theta})$ . For any fixed  $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$ ,  $g(\theta, \hat{\theta})$  is Lipschitz continuous in  $\theta$  by (a) and (b). Hence, it is absolutely continuous, and

$$g(\theta, \hat{\theta}) = \int_{\hat{\theta}}^{\theta} \frac{\partial g(q, \hat{\theta})}{\partial \theta} dq = \int_{\hat{\theta}}^{\theta} \left[ \bar{\Phi}'(q) - \frac{\partial \Phi(q, \hat{\theta})}{\partial \theta} \right] dq.$$

Therefore, for all  $\theta \in (\underline{\theta}, \bar{\theta})$ ,  $\Phi(\theta, \hat{\theta})$  is maximized by setting  $\hat{\theta} = \theta$  if and only if (9) holds. ■

In the special case of a static model, the necessity of the integral monotonicity condition (8) readily follows from Theorem 1 and Lemma 1: There, for any fixed message  $\hat{\theta}_i$ , agent  $i$ 's expected payoff can simply be assumed (equi-)Lipschitz continuous and differentiable in the true type  $\theta_i$ . By Theorem 1, this implies Lipschitz continuity of his equilibrium payoff in  $\theta_i$  so that necessity of integral monotonicity follows by Lemma 1.

In contrast, in the dynamic model, fixing agent  $i$ 's period- $t$  message  $\hat{\theta}_{it}$ , the Lipschitz continuity of his expected payoff in the current type  $\theta_{it}$  (or the formula for its derivative) cannot be assumed but must be derived from the agent's future optimizing behavior (see the counterexample in footnote

22). In particular, we show that in a Markov environment, the fact that the choice rule  $\langle \chi, \psi \rangle$  implementing  $\chi$  satisfies ICFOC (by Theorem 1) implies that the agent's expected payoff under a one-step deviation from truthtelling satisfies a condition analogous to ICFOC with respect to the modified choice rule  $\langle \chi \circ \hat{\theta}_{it}, \psi \circ \hat{\theta}_{it} \rangle$  induced by the lie. This step is non-trivial and uses the fact that, in a Markov environment, truthtelling is an optimal continuation strategy following the lie. Since the agent's expected equilibrium payoff at any truthful history is Lipschitz continuous in the current type by Theorem 1, the necessity of integral monotonicity then follows by Lemma 1.

The other key difference pertains to the sufficiency part of the result: In a static environment, the payments constructed using the envelope formula ICFOC guarantee that the agent's payoff under truthtelling is Lipschitz continuous and satisfies ICFOC by construction. Incentive compatibility then follows from integral monotonicity by Lemma 1. In contrast, in the dynamic model, the payments defined by (7) guarantee only that ICFOC is satisfied at truthful histories (by Theorem 2(i)). However, in a Markov environment it is irrelevant for the agent's continuation payoff whether he has been truthful in the past or not, and hence ICFOC extends to all histories. Thus, by Lemma 1, integral monotonicity implies that one-stage deviations from truthtelling are unprofitable. Establishing a one-stage-deviation principle for the environment then concludes the proof.<sup>27</sup>

**Remark 4** Theorem 3 can be extended to characterize *strong* PBE-implementability in non-Markov environments. However, for such environments, the restriction to strongly truthful PBE is in general with loss. In the supplementary material, we show that this approach nevertheless allows us to verify the implementability of the optimal allocation rule in some specific non-Markov environments.  $\square$

### 3.3.1 Verifying Integral Monotonicity

The integral monotonicity condition (8) is in general not an easy object to work with. This is true even in static models, except for the special class of environments where both the type and the allocation are unidimensional and the agent's payoff is supermodular, in which case integral monotonicity is equivalent to the monotonicity of the allocation rule. Our dynamic problem essentially never falls into this class: Even if the agent's current type is unidimensional, his report will in general affect the allocation both in the current period as well as in all future periods, which renders the allocation space multidimensional. For this reason, we provide monotonicity conditions which are stronger than integral monotonicity but easier to verify.<sup>28</sup> Some of these sufficient conditions apply only to environments satisfying additional restrictions:

<sup>27</sup>The usual version of the one-stage-deviation principle (for example, Fudenberg and Tirole, 1991, p.110) is not applicable since payoffs are a priori not continuous at infinity because flow payments need not be bounded.

<sup>28</sup>For similar sufficient conditions for static models with a unidimensional type and multidimensional allocation space, see, for example, Matthews and Moore (1987), whose condition is analogous to our strong monotonicity, and Garcia (2005), whose condition is analogous to our ex-post monotonicity.

**Condition (F-AUT) Process Autonomous:** For all  $i = 1, \dots, n$ ,  $t \geq 1$ , and  $\theta_i^{t-1} \in \Theta_i^{t-1}$ , the distribution  $F_{it}(\theta_i^{t-1}, x_i^{t-1})$  does not depend on  $x_i^{t-1}$ .

**Condition (F-FOSD) Process First-Order Stochastic Dominance:** For all  $i = 1, \dots, n$ ,  $t \geq 1$ ,  $\theta_{it} \in \Theta_{it}$ , and  $x_i^{t-1} \in X_i^{t-1}$ ,  $F_{it}(\theta_{it}|\theta_i^{t-1}, x_i^{t-1})$  is nonincreasing in  $\theta_i^{t-1}$ .

**Corollary 1 (Monotonicities)** Suppose the environment is regular and Markov. Then the integral monotonicity condition (8) of Theorem 3 is implied by any of the following conditions (listed in decreasing order of generality):

(i) **Single-crossing:** For all  $i = 1, \dots, n$ ,  $t \geq 0$ ,  $\theta^{t-1} \in \Theta^{t-1}$ ,  $\hat{\theta}_{it} \in \Theta_{it}$ , and a.e.  $\theta_{it} \in \Theta_{it}$ ,

$$\left[ D_{it}^{\chi, \Gamma}(\theta^{t-1}, \theta_{it}) - D_{it}^{\chi \circ \hat{\theta}_{it}, \Gamma}(\theta^{t-1}, \theta_{it}) \right] \cdot (\theta_{it} - \hat{\theta}_{it}) \geq 0.$$

(ii) **Average monotonicity:** For all  $i = 1, \dots, n$ ,  $t \geq 0$ , and  $(\theta^{t-1}, \theta_{it}) \in \Theta^{t-1} \times \Theta_{it}$ ,  $D_{it}^{\chi \circ \hat{\theta}_{it}, \Gamma}(\theta^{t-1}, \theta_{it})$  is nondecreasing in  $\hat{\theta}_{it}$ .

(iii) **Ex-post monotonicity:** Condition F-AUT holds, and for all  $i = 1, \dots, n$ ,  $t \geq 0$ , and  $\theta \in \Theta$ ,

$$\sum_{\tau=t}^{\infty} I_{i,(t),\tau}(\theta_i^\tau) \frac{\partial U_i(\theta, \chi(\hat{\theta}_{it}, \theta_{i,-t}, \theta_{-i}))}{\partial \theta_{i\tau}} \quad (10)$$

is nondecreasing in  $\hat{\theta}_{it}$ .

(iv) **Strong monotonicity:** Conditions F-AUT and F-FOSD hold, and for all  $i = 1, \dots, n$ ,  $t \geq 0$ , and  $\theta_{-i} \in \Theta_{-i}$ ,  $X_{it} \subseteq \mathbb{R}^m$ ,  $U_i(\theta, x)$  has increasing differences in  $(\theta_i, x_i)$  and is independent of  $x_{-i}$ , and  $\chi_i(\theta)$  is nondecreasing in  $\theta_i$ .

To see the relationship between the conditions, observe that the most stringent of the four, strong monotonicity, amounts to the requirement that each individual term in the sum in (10) be nondecreasing in  $\hat{\theta}_{it}$  (note that, under F-FOSD,  $I_{i,(t),\tau} \geq 0$ ). Ex-post monotonicity weakens this by requiring only that the sum be nondecreasing, which permits us to dispense with F-FOSD as well as with the assumption that  $U_i$  has increasing differences. By recalling the definition of the  $D$  functions in (6), we see that average monotonicity in turn weakens ex-post monotonicity by averaging over states, which also permits us to dispense with F-AUT. Finally, single-crossing relaxes average monotonicity by requiring that the expectation of the sum in (10) changes sign only once at  $\hat{\theta}_{it} = \theta_{it}$ , as opposed to being monotone in  $\hat{\theta}_{it}$ . But single-crossing clearly implies integral monotonicity, proving the corollary. The following example illustrates:

**Example 4** Consider a regular Markov environment with one agent whose allocation utility takes the form  $U(\theta, x) = \sum_{t=0}^{\infty} \delta^t \theta_t x_t$ , where for all  $t \geq 0$ , the period- $t$  consumption  $x_t$  is an element of some unidimensional set  $X_t \subseteq \mathbb{R}$ . Suppose that conditions F-AUT and F-FOSD hold. By (6),

$$D_t^{X \circ \hat{\theta}_t}(\theta^{t-1}, \theta_t) = \mathbb{E}^{\lambda|\theta^{t-1}, \theta_t} \left[ \sum_{\tau=t}^{\infty} \delta^\tau I_{(t), \tau}(\tilde{\theta}^\tau) \chi_\tau(\hat{\theta}_t, \tilde{\theta}_{-t}) \right].$$

Thus, average monotonicity requires that increasing the current message  $\hat{\theta}_t$  increases the agent's average discounted consumption, where period- $\tau$  consumption is discounted using the discount factor  $\delta$  as well as the impulse response  $I_{(t), \tau}(\tilde{\theta}^\tau)$  of period- $\tau$  signal to period- $t$  signal. Ex-post monotonicity requires that the discounted consumption  $\sum_{\tau=t}^{\infty} \delta^\tau I_{(t), \tau}(\theta^\tau) \chi_\tau(\hat{\theta}_t, \theta_{-t})$  be increasing in  $\hat{\theta}_t$  along every path  $\theta$ . And strong monotonicity requires that increasing  $\hat{\theta}_t$  increases consumption  $\chi_\tau(\hat{\theta}_t, \theta_{-t})$  in every period  $\tau \geq t$  irrespective of the agent's signals in the other periods.  $\square$

Courty and Li (2000) studied a two-period version of Example 4 with allocation  $x_1 \in X_1 = [0, 1]$  only in the second period (i.e.,  $X_t = \{0\}$  for all  $t \neq 1$ ), and provided sufficient conditions for implementability in two cases. The first (their Lemma 3.3) assumes F-FOSD and corresponds to our strong monotonicity. This case was extended to many agents by Eső and Szentes (2007).<sup>29</sup> The second case assumes that varying the initial signal  $\theta_0$  induces a mean-preserving spread by rotating  $F_1(\cdot|\theta_0)$  about a single point  $z$ . Courty and Li showed (as their Lemma 3.4) that it is then possible to implement any  $\chi_1$  that is non-increasing in  $\theta_0$ , non-decreasing in  $\theta_1$ , and satisfies “no under production:”  $\chi_1(\theta_0, \theta_1) = 1$  for all  $\theta_1 \geq z$ .<sup>30</sup> This case is covered by our ex-post monotonicity, which in period 0 requires that  $I_{(0), 1}(\theta_0, \theta_1) \chi_1(\hat{\theta}_0, \theta_1)$  be non-decreasing in  $\hat{\theta}_0$ . To see this, note that by the canonical impulse response formula (5), we have  $I_{(0), 1}(\theta_0, \theta_1) \leq 0$  (resp.,  $\geq 0$ ) if  $\theta_1 \leq z$  (resp.,  $\geq z$ ). Thus  $I_{(0), 1}(\theta_0, \theta_1) \chi_1(\hat{\theta}_0, \theta_1)$  is weakly increasing in  $\hat{\theta}_0$  if  $\theta_1 \leq z$ , because  $\chi_1$  is non-increasing in  $\theta_0$ , whereas it is constant in  $\hat{\theta}_0$  if  $\theta_1 \geq z$ , because  $\chi_1$  satisfies no under production.

The main application of Corollary 1 is in the design of optimal dynamic mechanisms, which we turn to in the next section. There, a candidate allocation rule is obtained by solving a suitable relaxed problem, and then Corollary 1 is used to verify that the allocation rule is indeed implementable. The results in the literature are typically based on strong monotonicity (for example, Battaglini, 2005, or Eső and Szentes, 2007) with the exception of the mean-preserving-spread case of Courty and Li (2000) discussed above. However, there are interesting applications where the optimal allocation rule fails to be strongly monotone, or where the kernels naturally depend on past decisions or fail first-order stochastic dominance. For instance, the optimal allocation rule in Example 5 below fails

<sup>29</sup>Eső and Szentes derived the result in terms of a state representation. Translated to the primitive types  $\theta_{i0}$  and  $\theta_{i1}$  (or  $v_i$  and  $V_i$  in their notation), their Corollary 1 shows that the allocation rule they are interested in implementing is strongly monotone. Note that Eső and Szentes's display (22) is a special case of our integral monotonicity condition, but stated in terms of a state representation. However, they used it only in conjunction with strong monotonicity.

<sup>30</sup>Courty and Li also considered the analogous case of “no over production,” to which similar comments apply.

strong monotonicity but satisfies ex-post monotonicity, whereas the optimal allocation rule in the bandit auctions of Section 5 fails ex-post monotonicity but satisfies average monotonicity.

**Remark 5** Suppose the environment is regular and Markov. Then for any allocation rule  $\chi$  that satisfies ex-post monotonicity, there exists a transfer rule  $\psi$  such that the complete information version of the model (where agents observe each others' types) has a subgame perfect Nash equilibrium in strongly truthful strategies. In other words, ex-post monotone allocation rules can be implemented in a periodic ex-post equilibrium in the sense of Athey and Segal (2013) and Bergemann and Välimäki (2010). This implies that any such rule can be implemented in a strongly truthful PBE of a direct mechanism where all reports, allocations, and payments are public. The transfers that guarantee this can be constructed as in (7) with the measures  $\lambda_i[\Gamma]|\theta^{t-1}, \theta_{it}$  replaced by the measure  $\lambda|\theta^t$ .  $\square$

## 4 Optimal Mechanisms

We now show how Theorems 1–3 can be used in the design of optimal dynamic mechanisms in Markov environments.<sup>31</sup> To this end, we introduce a principal (labeled as “agent 0”) whose payoff takes the quasilinear form  $U_0(\theta, x) - \sum_{i=1}^n \sum_{t=0}^{\infty} \delta^t p_{it}$  for some function  $U_0 : \Theta \times X \rightarrow \mathbb{R}$ . The principal seeks to design a PBIC mechanism to maximize his expected payoff. As is standard in the literature, we assume that the principal makes a take-it-or-leave-it offer to the agents in period zero, after each agent  $i$  has observed his initial type  $\theta_{i0}$ . Each agent can either accept the mechanism, or reject it to obtain his reservation payoff, which we normalize to 0 for all agents and types.<sup>32</sup> The principal’s mechanism design problem is thus to maximize his ex-ante expected payoff

$$\mathbb{E}^{\lambda[\chi]} \left[ U_0(\tilde{\theta}, \chi(\tilde{\theta})) - \sum_{i=1}^n \sum_{t=0}^{\infty} \delta^t \psi_{it}(\tilde{\theta}) \right] = \mathbb{E}^{\lambda[\chi]} \left[ \sum_{i=0}^n U_i(\tilde{\theta}, \chi(\tilde{\theta})) - \sum_{i=1}^n V_{i0}^{\langle \chi, \psi \rangle}(\tilde{\theta}_{i0}) \right]$$

by choosing a feasible choice rule  $\langle \chi, \psi \rangle$  that is PBIC and satisfies

$$V_{i0}^{\langle \chi, \psi \rangle}(\theta_{i0}) \geq 0 \quad \text{for all } i = 1, \dots, n \text{ and } \theta_{i0} \in \Theta_{i0}. \quad (11)$$

Any solution to this problem is referred to as an *optimal mechanism*.

<sup>31</sup>For other possible applications, see, for example, Skrzypacz and Toikka (2013) who consider the feasibility of efficient dynamic contracting in repeated trade and in other dynamic collective choice problems.

<sup>32</sup>If an agent can accept the mechanism, but can then quit at a later stage, participation constraints have to be introduced in all periods  $t \geq 0$ . However, in our quasilinear environment with unlimited transfers, the principal can ask the agent to post a sufficiently large bond upon acceptance, to be repaid later, so as to make it unprofitable to quit and forfeit the bond at any time during the mechanism. (With an infinite horizon, annuities can be used in place of bonds.) For this reason, we ignore participation constraints in periods  $t > 0$ . Note that in non-quasilinear settings where agents have a consumption-smoothing motive, bonding is costly, and hence participation constraints may bind in many periods (see, for example, Hendel and Lizzeri, 2003).



We restrict attention to regular environments throughout this section, and assume that the initial distribution  $F_{i0}$  of each agent  $i$  is absolutely continuous with density  $f_{i0}(\theta_{i0}) > 0$  for almost every  $\theta_{i0} \in \Theta_{i0}$ , and, for simplicity, that the set of initial types  $\Theta_{i0}$  is bounded from below (i.e.,  $\underline{\theta}_{i0} > -\infty$ ). Hence we can use ICFOC $_{i,0}$  (by Theorem 1) as in static settings to rewrite the principal's payoff as

$$\mathbb{E}^{\lambda[\chi]} \left[ \sum_{i=0}^n U_i(\tilde{\theta}, \chi(\tilde{\theta})) - \sum_{i=1}^n \frac{1}{\eta_{i0}(\tilde{\theta}_{i0})} \sum_{t=0}^{\infty} I_{i,(0),t}(\tilde{\theta}_i^t, \chi_i^{t-1}(\tilde{\theta})) \frac{\partial U_i(\tilde{\theta}, \chi(\tilde{\theta}))}{\partial \theta_{it}} \right] - \sum_{i=1}^n V_{i0}^{\langle \chi, \psi \rangle}(\underline{\theta}_{i0}),$$

where  $\eta_{i0}(\theta_{i0}) \equiv f_{i0}(\theta_{i1})/(1 - F_{i0}(\theta_{i1}))$  is the hazard rate of agent  $i$ 's period-0 type.<sup>33</sup> The first term above is the *expected (dynamic) virtual surplus*, which is only a function of the allocation rule  $\chi$ .

The principal's problem is in general analytically intractable. Hence we adopt the "first-order approach" typically followed in the literature. In particular, we consider a relaxed problem where PBIC is relaxed to the requirement that  $\langle \chi, \psi \rangle$  satisfy ICFOC $_{i,0}$  for all  $i$ , and where a participation constraint is imposed only on each agent's lowest initial type, that is, (11) is replaced with

$$V_{i0}^{\langle \chi, \psi \rangle}(\underline{\theta}_{i0}) \geq 0 \quad \text{for all } i = 1, \dots, n. \quad (12)$$

Since subtracting a constant from agent  $i$ 's period-0 transfer leaves ICFOC $_{i,0}$  unaffected but increases the principal's payoff, the constraints (12) must bind at a solution. It follows that an allocation rule  $\chi^*$  is part of a solution to our relaxed problem if, and only if,  $\chi^*$  maximizes

$$\mathbb{E}^{\lambda[\chi]} \left[ \sum_{i=0}^n U_i(\tilde{\theta}, \chi(\tilde{\theta})) - \sum_{i=1}^n \frac{1}{\eta_{i0}(\tilde{\theta}_{i0})} \sum_{t=0}^{\infty} I_{i,(0),t}(\tilde{\theta}_i^t, \chi_i^{t-1}(\tilde{\theta})) \frac{\partial U_i(\tilde{\theta}, \chi(\tilde{\theta}))}{\partial \theta_{it}} \right]. \quad (13)$$

This problem is in general a dynamic programming problem, and, in contrast to static settings, it cannot be solved pointwise in general. Note that as the definition of the relaxed problem uses ICFOC $_{i,0}$ , Theorem 1 plays a key role in identifying the candidate allocation rule  $\chi^*$ .

If the environment is Markov, then Theorem 3 can be used to verify whether the candidate allocation rule  $\chi^*$  is PBE-implementable (possibly by checking one of the conditions in Corollary 1). In case the answer is affirmative, Theorem 2 provides a formula for constructing a transfer rule  $\psi$  such that  $\langle \chi^*, \psi \rangle$  is PBIC. We can then subtract the constant  $V_{i0}^{\langle \chi^*, \psi \rangle}(\underline{\theta}_{i0})$  from each agent  $i$ 's initial transfer to get a PBIC choice rule  $\langle \chi^*, \psi^* \rangle$  such that  $V_{i0}^{\langle \chi^*, \psi^* \rangle}(\underline{\theta}_{i0}) = 0$  for all  $i = 1, \dots, n$ . Thus it remains to verify that all the other participation constraints in (11) are satisfied. As part of the next result, we show that F-FOSD is a sufficient condition for this provided that each agent's utility is increasing in his own type sequence (endowed with the pointwise order).

---

<sup>33</sup>Recall that the proof of Theorem 1 uses only deviations in which, in terms of a state representation, each agent  $i$  reports truthfully future shocks  $\varepsilon_{it}$ ,  $t > 0$ . Hence, as noted by Esó and Szentes (2007, 2013), this expression gives the principal's payoff also in a hypothetical environment where the shocks are observable to the principal. (This observation underlies the "irrelevance result" of Esó and Szentes, 2013.) However, the set of PBIC choice rules is strictly larger when the shocks are observable, so the principal's problems in the two settings are not equivalent in general.

**Corollary 2 (Optimal mechanisms)** *Suppose the environment is regular and Markov. Suppose in addition that Condition F-FOSD holds and, for all  $i = 1, \dots, n$  and  $(\theta, x) \in \Theta \times X$ ,  $U_i(\theta, x)$  is nondecreasing in  $\theta_i$ . Let  $\chi^*$  be an allocation rule that maximizes the expected virtual surplus (13), and suppose that, with some belief system  $\Gamma \in \mathbf{\Gamma}(\chi^*)$ , it satisfies the integral monotonicity condition (8) in all periods. Then the following statements hold:*

(i) *There exists a transfer rule  $\psi^*$  such that (a) the direct mechanism  $\langle \chi, \psi^* \rangle$  has a strongly truthful PBE with belief system  $\Gamma$  and where, for all  $i = 1, \dots, n$ , the flow payments  $\psi_{it}^*$ ,  $t \geq 0$ , can be disclosed to agent  $i$ ; (b) the period-0 participation constraints (11) are satisfied; and (c) the period-0 participation constraints of the lowest initial types (12) hold with equality.*

(ii) *The above choice rule  $\langle \chi^*, \psi^* \rangle$  maximizes the principal's expected payoff across all PBIC choice rules that satisfy participation constraints (11).*

(iii) *If  $\langle \chi, \psi \rangle$  is optimal for the principal among all PBIC choice rules that satisfy participation constraints (11), then  $\chi$  maximizes the expected virtual surplus (13).*

(iv) *The principal's expected payoff cannot be increased by using randomized mechanisms.*

**Remark 6** The statements (ii)-(iv) remain true if PBIC is weakened to the requirement that there exists a Bayesian-Nash equilibrium in on-path truthful strategies. This is because the derivation of the expected virtual surplus (13) uses only  $\text{ICFOC}_{i,0}$ , which by Remark 1 holds under this weaker notion of incentive compatibility.  $\square$

**Proof.** Parts (i)(a) and (i)(c) follow by the arguments preceding the corollary. For (i)(b), note that under F-FOSD, impulse responses are non-negative almost everywhere, and hence each  $V_{i0}^{\langle \chi^*, \psi^* \rangle}(\theta_{i0})$  is nondecreasing in  $\theta_{i0}$  by the envelope formula (1) given that  $U_i$  is nondecreasing in  $\theta_i$ .

Parts (ii) and (iii) follow by the arguments preceding the corollary.

Finally, for part (iv) note that a randomized mechanism is equivalent to a mechanism that conditions on the random types of a fictitious agent. Since the expected virtual surplus in this augmented setting is independent of the signals of the fictitious agent and still takes the form (13), it is still maximized by the non-randomized allocation rule  $\chi^*$ . Thus, applying parts (i) and (ii) to the augmented setting implies that the deterministic choice rule  $\langle \chi^*, \psi^* \rangle$  maximizes the principal's expected payoff. (A similar point was made by Strausz (2006) for static mechanisms.)  $\blacksquare$

Note that F-FOSD and the assumption that the agents' utilities be nondecreasing in own type are only used to establish that each agent  $i$ 's equilibrium payoff  $V_{i0}^{\langle \chi^*, \psi^* \rangle}(\theta_{i0})$  is minimized at  $\underline{\theta}_{i0}$ . If this conclusion can be arrived at by some other means (for example, by using (1) to solve for the function  $V_{i0}^{\langle \chi^*, \psi^* \rangle}$ ), these assumptions can be dispensed with.

Corollary 2 provides a guess-and-verify approach analogous to that typically followed in static settings. We illustrate its usefulness below by using it to discuss optimal distortions in dynamic contracts, and to find optimal “bandit auctions.” Similarly to static settings, however, the conditions

under which the relaxed problem has an implementable solution are by no means generic. As pointed out by Battaglini and Lamba (2012), a particularly problematic case obtains when the type of an agent remains constant with high probability, but nevertheless has a small probability of being renewed. In terms of our analysis, the problem is that then the impulse response becomes highly non-monotone in the current type, which in turn may result in the allocation being so non-monotone in the current type that integral monotonicity is violated.<sup>34</sup>

It is, of course, possible to reverse-engineer conditions that guarantee that the relaxed problem has an implementable solution, but given the complexity of the problem, such conditions tend to be grossly sufficient. Nevertheless, for completeness we provide sufficient conditions for an allocation rule that maximizes expected virtual surplus to satisfy strong monotonicity of Corollary 1.

**Condition (U-COMP) *Utility Complementarity*:**  $X$  is a lattice,<sup>35</sup> and for all  $i = 0, \dots, n$ ,  $t \geq 0$  and  $\theta \in \Theta$ ,  $U_i(\theta, x)$  is supermodular in  $x$ , and  $-\partial U_i(\theta, x) / \partial \theta_{it}$  is supermodular in  $x$ .

This condition holds weakly in the special case where  $X_t$  is a subset of  $\mathbb{R}$  in every period  $t$ , and the payoffs  $U_i(\theta, x)$  are additively separable in  $x_t$ . More generally, U-COMP allows for strict complementarity across time, e.g., as in habit-formation models where higher consumption today increases the marginal utility of consumption tomorrow. On the other hand, U-COMP is not satisfied when allocating private goods in limited supply as in auctions.

**Condition (U-DSEP) *Utility Decision-Separable*:**  $X = \prod_{t=0}^{\infty} X_t$  and, for all  $i = 0, \dots, n$ , and  $(\theta, x) \in \Theta \times X$ ,  $U_i(\theta, x) = \sum_{t=0}^{\infty} \delta^t u_{it}(\theta^t, x_t)$ .

**Proposition 1 (Primitive conditions for strong monotonicity)** *Suppose the environment is regular and Markov, Conditions F-AUT and F-FOSD hold, and for all  $i = 0, \dots, n$ , and  $t \geq 0$ ,  $X_{it}$  is a subset of an Euclidean space. Suppose that either of the following conditions is satisfied:*

(i) *Condition U-COMP holds, and for all  $i = 1, \dots, n$ , agent  $i$ 's virtual utility*

$$U_i(\theta, x) - \frac{1}{\eta_{i0}(\theta_{i0})} \sum_{t=0}^{\infty} \delta^t I_{i,(0),t}(\theta_i^t) \frac{\partial u_i(\theta_t, x^t)}{\partial \theta_{it}}$$

*has increasing differences in  $(\theta, x)$ , and the same is true of the principal's utility  $U_0(\theta, x)$ .*

(ii) *Condition U-DSEP holds, and for all  $i = 1, \dots, n$ , and  $t \geq 0$ ,  $X_{it} \subseteq \mathbb{R}$  and there exists a*

---

<sup>34</sup>For a concrete example, consider a single-agent environment with  $\tilde{\theta}_0$  distributed uniformly on  $[0, 1]$ . Suppose that  $\tilde{\theta}_1 = \tilde{\theta}_0$  with probability  $q$ , and that with the complementary probability  $\tilde{\theta}_1$  is drawn uniformly from  $[0, 1]$  independently of  $\tilde{\theta}_0$ . The period-0 impulse response is then  $I_{(0),1}(\theta_0, \theta_1) = \mathbf{1}_{\{\theta_0 = \theta_1\}}$ . By inspection of the expected virtual surplus (13), the period-1 allocation is thus distorted only if  $\theta_1 = \theta_0$ .

<sup>35</sup>The assumption that  $X$  is a lattice is not innocuous when  $n > 1$ : For example, it holds when each  $x_t$  describes the provision of a one-dimensional public good, but it need not hold if  $x_t$  describes the allocation of a private good.

nondecreasing function  $\varphi_{it} : \Theta_i^t \rightarrow \mathbb{R}^m$ , with  $m \leq t$ , such that agent  $i$ 's virtual flow utility

$$u_{it}(\theta_t, x_t) - \frac{1}{\eta_{i0}(\theta_{i0})} I_{i,(0),t}(\theta_i^t) \frac{\partial u_{it}(\theta_t, x_t)}{\partial \theta_{it}}$$

depends only on  $\varphi_{it}(\theta_i^t)$  and  $x_{it}$ , and has strictly increasing differences in  $(\varphi_{it}(\theta_i^t), x_{it})$ , while the principal's flow utility depends only on  $x_t$ .

Then, if the problem of maximizing expected virtual surplus (13) has a solution, it has a solution  $\chi$  such that, for all  $i = 1, \dots, n$  and  $\theta_{-i} \in \Theta_{-i}$ ,  $\chi_i(\theta_i, \theta_{-i})$  is nondecreasing in  $\theta_i$ .

When F-AUT and U-DSEP hold (i.e., types evolve independently of decisions and payoffs are separable in decisions), expected virtual surplus (13) can be maximized pointwise, which explains why condition (ii) of Proposition 1 only involves flow payoffs. Special cases of this result appear in Courty and Li (2000), who provided sufficient conditions for strong monotonicity by means of parametric examples, and in Esó and Szentes (2007), whose Assumptions 1 and 2 imply that  $I_{i,(0),1}(\theta_{i0}, \theta_{i1})$  is nonincreasing in both  $\theta_{i0}$  and  $\theta_{i1}$ , which together with their payoff functions imply condition (ii) (with  $\varphi_{it} = \text{id}$ ). For a novel setting that satisfies condition (ii), see Example 6 below.

By inspection of Proposition 1, guaranteeing strong monotonicity requires single-crossing and third-derivative assumptions familiar from static models. The new assumptions that go beyond them concern the impulse response functions. This is best illustrated by considering even stronger sufficient conditions, which can be stated separately on utilities and processes. For concreteness, suppose that U-DSEP holds and  $X_t$  is one-dimensional (so that either case in the proposition can be applied). Then, in the initial period  $t = 0$ , it suffices to impose the static conditions: for each agent  $i$ , the allocation utility  $u_{i0}(\theta_0, x_0)$  and the partial  $-\partial u_{i0}(\theta_0, x_0) / \partial \theta_{i0}$  have increasing differences (ID) in allocation and types (the latter being a third-derivative condition on the utility function), and the hazard rate  $\eta_{i0}(\theta_{i0})$  is nondecreasing. In periods  $t \geq 1$ , in addition to imposing the static conditions to current utility flows, it suffices to assume that the impulse response  $I_{i\tau}(\theta_i^\tau)$  be nondecreasing in types. This implies that the term capturing the agent's information rent,  $-\frac{1}{\eta_{i0}(\theta_{i0})} I_{it}(\theta_i^t) \partial u_{it}(\theta^t, x_t) / \partial \theta_{it}$  has ID in allocation and types. Heuristically, nondecreasing impulse responses lead to distortions being decreasing in types, which helps to ensure monotonicity of the allocation.

**Remark 7** We discuss in the supplementary material how Corollary 2 and Proposition 1 can be adapted to finding optimal mechanisms in some non-Markov environments. Note that the derivation of the expected virtual surplus (13) above makes no reference to Markov environments, and hence the difference is in verifying that the allocation rule maximizing it is indeed implementable.  $\square$

## 4.1 Distortions

A first-best allocation rule maximizes the expected surplus  $\mathbb{E}^{\lambda[\chi]} \left[ \sum_{i=0}^n U_i(\tilde{\theta}, \chi(\tilde{\theta})) \right]$  in our quasilinear environment. Similarly to static setting, a profit maximizing principal introduces distortions to the

allocations to reduce the agents' expected information rents. When the participation constraints of the lowest initial types in (12) bind, the expected rent of agent  $i$  is given by

$$\mathbb{E}^{\lambda[\chi]} \left[ \sum_{t=0}^{\infty} \frac{1}{\eta_{i0}(\tilde{\theta}_{i0})} I_{i,(0),t}(\tilde{\theta}_i^t, \chi_i^{t-1}(\tilde{\theta})) \frac{\partial U_i(\tilde{\theta}, \chi(\tilde{\theta}))}{\partial \theta_{it}} \right].$$

Thus the period-0 impulse response functions are an important determinant of the rent, and, by implication, of the distortions in optimal dynamic mechanisms. Given the various forms these functions may take, little can be said about the nature of these distortions in general. Indeed, we illustrate by means of a simple class of single-agent environments that the distortion in period  $t$  may be a non-monotone function of the agent's types, or, for a fixed type sequence, a non-monotone function of the time period  $t$ . Example 5 also illustrates the use of ex-post monotonicity to verify implementability.

**Example 5 (Nonlinear AR process)** Consider a buyer-seller relationship, which lasts for  $T + 1$  periods, with  $T \leq \infty$ . The buyer's payoff takes the form  $U_1(\theta, x) = \sum_{t=0}^T \delta^t (a + \theta_t) x_t$ , with  $a \in \mathbb{R}_{++}$  and  $X_t = [0, \bar{x}]$  for some  $\bar{x} \gg 0$ . The seller's payoff is given by  $U_0(\theta, x) = -\sum_{t=0}^T \delta^t \frac{x_t^2}{2}$ . The buyer's type evolves according to the *nonlinear* AR process  $\theta_t = \phi(\theta_{t-1}) + \varepsilon_t$ , where  $\phi$  is an increasing differentiable function, with  $\phi(0) = 0$ ,  $\phi(1) < 1$ , and  $\phi' \leq b$  for some  $1 \leq b < \frac{1}{\delta}$ , and where the shocks  $\varepsilon_t$  are independent over time, with support  $[0, 1 - \phi(1)]$ . By putting  $z_t(\theta_{t-1}, \varepsilon_t) = \phi(\theta_{t-1}) + \varepsilon_t$  and using formula (4), we find the period-0 impulse responses  $I_{(0),t}(\theta^t) = \prod_{\tau=0}^{t-1} \phi'(\theta_\tau)$ .

Since the type process is autonomous and decisions are separable across time (i.e., F-AUT and U-DSEP hold), the first-best allocation rule simply sets  $x_t = a + \theta_t$  for all  $t \geq 0$  and  $\theta^t \in \Theta^t$ . Furthermore, we can maximize expected virtual surplus (13) pointwise to find the allocation rule

$$\chi_t(\theta^t) = \max \left\{ 0, a + \theta_t - \frac{1}{\eta_0(\theta_0)} \prod_{\tau=0}^{t-1} \phi'(\theta_\tau) \right\} \text{ for all } t \geq 0 \text{ and } \theta^t \in \Theta^t.$$

We show in the supplementary material that if the hazard rate  $\eta_0$  is nondecreasing, then  $\chi$  is ex-post monotone, and thus it is an optimal allocation rule by Corollaries 1 and 2. Note that  $\chi$  exhibits downward distortions since  $\phi' > 0$ . Increasing the period- $\tau$  type  $\theta_\tau$  for  $1 \leq \tau < t$ , reduces distortions in period  $t$  if  $\phi$  is concave, but increases distortions if  $\phi$  is convex. (Note that in the latter case,  $\chi$  is not strongly monotone, yet it is PBE-implementable.) When  $\phi$  is neither concave nor convex, the effect is non-monotone. Similarly, if  $\phi'(\theta_{t-1}) < 1$ , then the distortion in period  $t$  is smaller than that in period  $t - 1$ , whereas if  $\phi'(\theta_{t-1}) > 1$ , then the period- $t$  distortion exceeds that in period  $t - 1$ .

Finally, note that the period- $t$  allocation  $\chi_t(\theta^t)$  is in general a non-trivial function of the buyer's types in all periods  $0, \dots, t$ . This is in contrast to the special case of a linear function  $\phi(\theta_t) = \gamma\theta_t$ ,  $\gamma > 0$ , considered by Besanko (1985), where the impulse response is the time-varying scalar  $I_t = \gamma^t$  as in Example 2, and where  $\chi_t(\theta^t)$  depends only on the initial type  $\theta_0$  and the current type  $\theta_t$ .  $\square$

The distortions in Example 5 are independent of the agent's current report. However, it is easy to construct examples where distortions are non-monotone also with respect to the current report:

**Example 6** Consider the environment of Example 5, but assume now that  $T = 1$  and that the buyer's type evolves as follows: The initial type  $\tilde{\theta}_0$  is distributed uniformly on  $\Theta_0 = [0, 1]$ , whereas  $\tilde{\theta}_1$  is distributed on  $\Theta_1 = [0, 1]$  according to the c.d.f.  $F_1(\theta_1|\theta_0) = \theta_1 - 2(\theta_0 - \frac{1}{2})\theta_1(1 - \theta_1)$  with linear density  $f_1(\theta_1|\theta_0) = 1 - 2(\theta_0 - \frac{1}{2})(1 - 2\theta_1)$  strictly positive on  $\Theta_1$  for all  $\theta_0$ . For  $\theta_0 = 1/2$ ,  $\tilde{\theta}_1$  is distributed uniform on  $[0, 1]$ . For  $\theta_0 < (>)1/2$  the density slopes downwards (upwards). Note that  $F$  satisfies F-FOSD. The canonical impulse response formula (5) from Example 3 gives

$$I_{(0),1}(\theta_0, \theta_1) = \frac{2\theta_1(1 - \theta_1)}{1 - 2(\theta_0 - \frac{1}{2})(1 - 2\theta_1)}.$$

The allocation rule that solves the relaxed program is then given by

$$\chi_0(\theta_0) = \max\{0, a + \theta_0 - (1 - \theta_0)\}, \quad \chi_1(\theta^1) = \max\left\{0, a + \theta_1 - (1 - \theta_0) \frac{2\theta_1(1 - \theta_1)}{1 - 2(\theta_0 - \frac{1}{2})(1 - 2\theta_1)}\right\}.$$

Because  $\chi$  is strongly monotone, it is clearly implementable. By inspection, the second-period allocation is efficient at the extremes, i.e., for  $\theta_1 \in \{0, 1\}$ , whereas all interior types are distorted downwards. Note that the no-distortions-at-the-bottom result is non-trivial, since even the lowest type here consumes a positive amount, so there would be room to distort downwards.  $\square$

In the supplementary material, we use monotone comparative statics to give sufficient conditions for the allocation rule maximizing expected virtual surplus (13) to exhibit downward distortions as in the above examples. However, upward distortions can naturally arise in applications. This was first shown by Courty and Li (2000), who provided a two-period example where the distribution of the agent's second-period type is ordered by his initial signal in the sense of a mean-preserving spread.

In consumption problems such as Example 5 and the one studied by Courty and Li (2000), distortions can be understood purely in terms of the canonical impulse response (see Example 1)

$$I_t(\theta^t) = \prod_{\tau=1}^t \left( - \frac{\partial F_\tau(\theta_\tau|\theta_{\tau-1})/\partial\theta_{\tau-1}}{f_\tau(\theta_\tau|\theta_{\tau-1})} \right).$$

If the kernels satisfy F-FOSD, then  $I_t(\theta^t)$  is positive leading to downward distortions as in Example 5. If F-FOSD fails, then  $I_t(\theta^t)$  is negative at some  $\theta^t$  yielding upward distortions at that history as in Courty and Li (2000). Dynamics can be seen similarly: As in Example 5, an increase in the impulse response increases distortions compared to the previous period, whereas a decrease leads to consumption being more efficient. In particular, if  $I_t(\theta^t) \rightarrow 0$ , then consumption converges to the first best over time.

## 5 Bandit Auctions

To illustrate our results, we consider the problem of a profit-maximizing seller who must design a sequence of auctions to sell off, in each period  $t \geq 0$ , an indivisible, non-storable good to a set of  $n \geq 1$  bidders who update their valuations upon consumption, i.e., upon winning the auction. This setting captures novel applications such as repeated sponsored search auctions where the advertisers privately learn about the profitability of clicks on their ads, or repeated procurement with learning-by-doing. It provides a natural environment where the kernels depend on past allocations.

Let  $X_{it} \in \{0, 1\}$  for all  $i = 0, \dots, n$  and  $t \geq 0$ , and define the set of feasible allocation sequences by  $X = \{x \in \prod_{t=0}^{\infty} \prod_{i=0}^n X_{it} : \sum_{i=0}^n x_{it} = 1 \text{ for all } t \geq 0\}$ . The seller's payoff function is then given by  $U_0(\theta, x) = -\sum_{t=0}^{\infty} \delta^t \sum_{i=1}^n x_{it} c_{it}$ , where  $c_{it} \in \mathbb{R}$  is the cost of allocating the object to bidder  $i$  (with  $c_{0t}$  normalized to 0). Each bidder  $i$ 's payoff function takes the form  $U_i(\theta, x) = \sum_{t=0}^{\infty} \delta^t \theta_{it} x_{it}$ .

The type process of bidder  $i = 1, \dots, n$  is constructed as follows. Let  $R_i = (R_i(\cdot|k))_{k \in \mathbb{N}}$  be a sequence of absolutely continuous, strictly increasing c.d.f.'s with mean bounded in absolute value uniformly in  $k$ . The first-period valuation  $\theta_{i0}$  is drawn from  $\Theta_{i0}$ , with  $\underline{\theta}_{i0} > -\infty$ , according to an absolutely continuous, strictly increasing c.d.f.  $F_{i0}$ . For all  $t > 0$ ,  $\theta_i^t \in \Theta_i^t$ , and  $x_i^{t-1} \in X_i^{t-1}$ , if  $x_{i,t-1} = 1$ , then

$$F_{it}(\theta_{it} | \theta_{i,t-1}, x_i^{t-1}) = R_i(\theta_{it} - \theta_{i,t-1} | \sum_{\tau=0}^{t-1} x_{i\tau});$$

if, instead,  $x_{i,t-1} = 0$ , then

$$F_{it}(\theta_{it} | \theta_{i,t-1}, x_i^{t-1}) = \begin{cases} 0 & \text{if } \theta_{it} < \theta_{i,t-1}, \\ 1 & \text{if } \theta_{it} \geq \theta_{i,t-1}. \end{cases}$$

This formulation embodies the following key assumptions: (1) Bidders' valuations change only upon winning the auction (i.e., if  $x_{it} = 0$ , then  $\theta_{i,t+1} = \theta_{it}$  almost surely); (2) The valuation processes are time-homogenous (i.e., if bidder  $i$  wins the object in period  $t$ , then the distribution of his period- $t+1$  valuation depends only on his period- $t$  valuation and the total number of times he won in the past).<sup>36</sup>

We start by verifying that the bandit auction environment defined above is regular and Markov. Each bidder  $i$ 's payoff function  $U_i$  clearly satisfies Conditions U-D, U-ELC, and U-SPR since each

---

<sup>36</sup>This kind of structure arises, for example, in a Bayesian learning model with Gaussian signals. That is, suppose each bidder  $i$  has a constant but unknown true valuation  $v_i$  for the object and starts with a prior belief  $v_i \sim N(\theta_{i0}, \tau_i)$  where precision  $\tau_i$  is common knowledge. Bidder  $i$ 's initial type  $\theta_{i0}$  is the mean of the prior distribution, which can have any distribution  $F_{i0}$  bounded from below. Each time upon winning the auction, bidder  $i$  receives a conditionally i.i.d. private signal  $s_i \sim N(v_i, \sigma_i)$  and updates his expectation of  $v_i$  using standard projection formulae. Take  $\theta_{it}$  to be bidder  $i$ 's posterior expectation in period  $t$ . Then  $R_i(\cdot|k)$  is the c.d.f. for the change in the posterior expectation due to the  $k$ -th signal, which is indeed independent of the current value of  $\theta_{it}$ . (Standard calculations show that  $R_i(\cdot|k)$  is in fact a Normal distribution with mean zero and variance decreasing in  $k$ .) Alternative specifications for  $R_i$  can be used to model learning-by-doing, habit formation, etc.

$X_{it}$  is bounded. As for bidder  $i$ 's type process, it satisfies the bounds in part (ii) of the Markov definition (and thus F-BE) by the uniform bound on the means of  $(R_i(\cdot|k))_{k \in \mathbb{N}}$ . In order to verify F-BIR, we use the canonical representation of the process. That is, for all  $t > 0$ ,  $\varepsilon_{it} \in (0, 1)$ , and  $(\theta_{i,t-1}, x_i^{t-1}) \in \Theta_{i,t-1} \times X_i^{t-1}$ , let

$$z_{it}(\theta_i^{t-1}, x_i^{t-1}, \varepsilon_{it}) = F_{it}^{-1}(\varepsilon_{it} | \theta_{i,t-1}, x_i^{t-1}) = \theta_{i,t-1} + \mathbf{1}_{\{x_{i,t-1}=1\}} R_i^{-1}(\varepsilon_{it} | \sum_{\tau=0}^{t-1} x_{i\tau}).$$

The  $Z$  functions then take the form

$$Z_{i,(s),t}(\theta_i^s, x_i^{t-1}, \varepsilon_i^t) = \theta_{is} + \sum_{m=s+1}^t \mathbf{1}_{\{x_{i,m-1}=1\}} R_i^{-1}(\varepsilon_{im} | \sum_{\tau=0}^{m-1} x_{i\tau}), \quad (14)$$

and hence  $\frac{\partial Z_{i,(s),t}}{\partial \theta_{is}} = 1$ . Therefore, F-BIR holds, and the impulse responses satisfy  $I_{i,(s),t}(\theta_i^t, x_i^{t-1}) = 1$  for all  $i = 1, \dots, n$ ,  $t \geq s \geq 0$ ,  $\theta_i^t \in \Theta_i^t$ , and  $x_i^{t-1} \in X_i^{t-1}$ .

Since the impulse responses  $I_{i,(0),t}$  are identical to 1 for all agents  $i$  and all periods  $t$ , the envelope formula (1) takes the form  $dV_{i0}^{(\chi, \psi)}(\theta_{i0})/d\theta_{i0} = \mathbb{E}^{\lambda_i[\chi]|\theta_{i0}} \left[ \sum_{t=0}^{\infty} \delta^t \chi_{it}(\tilde{\theta}) \right]$ , and the problem of maximizing expected virtual surplus (13) becomes

$$\sup_{\chi \in \mathcal{X}} \mathbb{E}^{\lambda[\chi]} \left[ \sum_{t=0}^{\infty} \delta^t \sum_{i=1}^n \left( \tilde{\theta}_{it} - c_{it} - \frac{1}{\eta_{i0}(\tilde{\theta}_{i0})} \right) \chi_{it}(\tilde{\theta}^t) \right].$$

This is a standard multi-armed bandit problem: The safe arm corresponds to the seller and yields a sure payoff equal to 0; the risky arm  $i = 1, \dots, n$  corresponds to bidder  $i$  and yields a flow payoff  $\theta_{it} - c_{it} - [\eta_{i0}(\theta_{i0})]^{-1}$ . The solution takes the form of an index policy. That is, define the virtual index of bidder  $i = 1, \dots, n$  in period  $t \geq 0$  given history  $(\theta_i^t, x_i^{t-1}) \in \Theta_i^t \times X_i^{t-1}$  as

$$\gamma_{it}(\theta_i^t, x_i^{t-1}) \equiv \max_T \mathbb{E}^{\lambda[\bar{\chi}_i]|\theta_{it}, x_i^{t-1}} \left[ \frac{\sum_{\tau=t}^T \delta^\tau \left( \tilde{\theta}_{i\tau} - c_{i\tau} - \frac{1}{\eta_{i0}(\tilde{\theta}_{i0})} \right)}{\sum_{\tau=t}^T \delta^\tau} \right], \quad (15)$$

where  $T$  is a stopping time, and  $\bar{\chi}_i$  is the allocation rule that assigns the object to bidder  $i$  in all periods. (Note that the virtual index depends on  $x_i^{t-1}$  only through  $\sum_{\tau=0}^{t-1} x_{i\tau}$ .) The index of the seller is identically equal to zero and for convenience we write it as  $\gamma_{0t}(\theta_0^t, x_0^{t-1}) \equiv 0$ . The following *virtual index policy* then maximizes the expected virtual surplus:<sup>37</sup> For all  $i = 1, \dots, n$ ,  $t \geq 0$ ,  $\theta^t \in \Theta^t$ , and  $x^{t-1} \in X^{t-1}$ , let  $J(\theta^t, x^{t-1}) \equiv \arg \max_{j \in \{0, \dots, n\}} \gamma_{jt}(\theta_j^t, x_j^{t-1})$ , and let

$$\chi_{it}(\theta^t) = \begin{cases} 1 & \text{if } i = \min J(\theta^t, x^{t-1}), \\ 0 & \text{otherwise.} \end{cases} \quad (16)$$

**Proposition 2 (Optimal bandit auctions)** *Suppose that for all  $i = 1, \dots, n$ , the hazard rate  $\eta_{i0}(\theta_{i0})$  is nondecreasing in  $\theta_{i0}$ . Let  $\langle \chi, \psi \rangle$  be the choice rule where  $\chi$  is the virtual index policy*

<sup>37</sup>The optimality of index policies is well known (e.g., Whittle, 1982 or Bergemann and Välimäki, 2008).



defined by (16) and where  $\psi$  is the transfer rule defined by (7). Then  $\langle \chi, \psi \rangle$  is an optimal mechanism in the bandit auction environment.

The environment is regular and Markov, F-FOSD holds, and each  $U_i$  is nondecreasing in  $\theta_i$ . Hence the result follows from Corollary 2 once we show that the virtual index policy  $\chi$  satisfies integral monotonicity. We do this in the appendix by showing that  $\chi$  satisfies average monotonicity defined in Corollary 1, which here requires that, for all  $i = 1, \dots, n$ ,  $s \geq 0$ , and  $(\theta^{s-1}, \theta_{is}) \in \Theta^{s-1} \times \Theta_{is}$ , bidder  $i$ 's expected discounted consumption

$$\mathbb{E}^{\lambda_i[\chi \circ \hat{\theta}_{is}, \Gamma] | \theta^{s-1}, \theta_{is}} \left[ \sum_{t=s}^{\infty} \delta^t (\chi \circ \hat{\theta}_{is})_{it}(\tilde{\theta}) \right]$$

is non-decreasing in his current bid  $\hat{\theta}_{is}$ . Heuristically, this follows because a higher bid in period  $s$  increases the virtual index of arm  $i$ , which results in bidder  $i$  consuming sooner in the sense that, for any  $k \in \mathbb{N}$ , the expected waiting time until he wins the auction for the  $k$ th time after period  $s$  is then weakly shorter. Note that because of learning, averaging is important: Even if increasing the current bid always makes bidder  $i$  to be more likely to win the auction today, for bad realizations of the resulting new valuation it leads to a lower chance of winning the auction in the future. However, by F-FOSD, higher current types are also more likely to win in the future on average.

It is instructive to compare the virtual index policy from the optimal bandit auction to the first-best index policy that maximizes social surplus. The first-best policy is implementable by using the team mechanism of Athey and Segal (2013), or the dynamic pivot mechanism of Bergemann and Välimäki (2010) who consider a similar bandit setting as an application. In the first-best policy, bidder  $i$ 's index at period- $t$  history  $(\theta_i^t, x_i^{t-1})$  is given by

$$g_{it}(\theta_i^t, x_i^{t-1}) \equiv \max_T \mathbb{E}^{\lambda[\bar{x}_i] | \theta_{it}, x_i^{t-1}} \left[ \frac{\sum_{\tau=t}^T \delta^\tau (\tilde{\theta}_{i\tau} - c_{it})}{\sum_{\tau=t}^T \delta^\tau} \right].$$

By inspection of (15) we see that the virtual index  $\gamma_{it}(\theta_i^t, x_i^{t-1})$  differs from the first-best index  $g_{it}(\theta_i^t, x_i^{t-1})$  only for the presence of the term  $\frac{1}{\eta_{i0}(\theta_{i0})}$ , which can be interpreted as bidder  $i$ 's “handicap.” In particular, note that the handicaps are determined by the bidders’ first-period (reported) types. Thus the optimal mechanism can be implemented by using the bidders’ initial reports to determine their handicaps along with the period-0 allocation and transfers, and by then running a handicapped efficient mechanism in periods  $t > 0$ , where the indices are computed as if the seller’s cost of assigning the good to bidder  $i$  was  $c_{it} - \frac{1}{\eta_{i0}(\theta_{i0})}$ .<sup>38</sup> This implies that even ex ante symmetric bidders will in general be treated asymmetrically in the future, and hence the distortions in future periods reflect findings in optimal static auctions with asymmetric bidders. (For example, the first-

<sup>38</sup>Board (2007) and Esó and Szentes (2007) find similar optimal mechanisms in settings where the type processes are autonomous and there is only one good to be allocated.

best and virtual indices will sometimes disagree on the ranking of any given bidders  $i$  and  $j$ , and hence  $i$  may win the object in some period  $t$  even if the first-best policy would award it to  $j$ .)

We conclude that the optimal mechanism for selling experience goods is essentially a dynamic auction with memory that grants preferential treatment based on the bidders' initial types. These features are markedly different from running a sequence of second-price auctions with a reserve price, and suggest potential advantages of building long-term contractual relationships in repeated procurement and sponsored search.

**Remark 8** Subsequent to the first version of our manuscript, Kakade et al (2011) considered a class of allocation problems that generalize our bandit auction environment, and showed that the optimal mechanism is a virtual version of the dynamic pivot mechanism of Bergemann and Välimäki (2010), the handicap mechanism being a special case. Postulating the model in terms of a state representation, they derived the allocation rule using our first-order approach, and established incentive compatibility in period 0 by verifying a condition analogous to our average monotonicity.

Kakade et al's proof of incentive compatibility for periods  $t > 0$  differs from ours, and relies on the above observation about the optimal mechanism from period 1 onwards being an efficient mechanism for a fictitious environment where the seller's cost of assigning the object to bidder  $i$  is  $c_{it} - \frac{1}{\eta_{i0}(\hat{\theta}_{i0})}$ , where  $\hat{\theta}_{i0}$  is  $i$ 's initial *report*. In particular, using an efficient mechanism for this fictitious environment that asks the bidders to re-report their initial types in period 1 gives the existence of a truthful continuation equilibrium from period 1 onwards. This approach requires, however, that every agent  $i$ 's payoff and state representation be separable in the sense that there exist functions  $\alpha_i$ ,  $\gamma_{it}$ , and  $\beta_{it}$ ,  $t \geq 0$ , such that (i)  $u_{it}(Z_{i,(0)}^t(\theta_{i0}, \varepsilon_i), x^t) = \alpha_i(\theta_{i0})\gamma_{it}(x^t) + \beta_{it}(\varepsilon_i^t, x^t)$  for all  $t$ , or that (ii)  $u_{it}(Z_{i,(0)}^t(\theta_{i0}, \varepsilon_i), x^t) = \alpha_i(\theta_{i0})\beta_{it}(\varepsilon_i^t, x^t)$  for all  $t$ . While our bandit auction environment satisfies condition (i) by inspection of (14), neither condition is satisfied in non-linear environments such as our Example 5, for which our approach of verifying integral monotonicity in every period is applicable. On the other hand, Kakade et al can accommodate non-Markov environments that are separable in the above sense. Thus the approaches are best viewed as complementary.  $\square$

## 6 Concluding Remarks

We extend the standard Myersonian approach to mechanism design to dynamic quasilinear environments. Our main results characterize local incentive compatibility constraints, provide a method of constructing transfers to satisfy them, address the uniqueness of these transfers, and give necessary and sufficient conditions for the implementability of allocation rules in Markov environments. These results lend themselves to the design of optimal dynamic mechanisms along the familiar lines of finding an allocation rule by maximizing expected (dynamic) virtual surplus, and then verifying that the allocation rule is implementable by checking appropriate monotonicity conditions.

The analysis permits a unified view of the existing literature by identifying general principles and highlighting what drives similarities and differences in the special cases considered. The generality of our model offers flexibility that facilitates novel applications, such as the design of sales mechanisms for the provision of new experience goods, or “bandit auctions.”

Our limited use of a state representation, also known as the *independent-shocks (IS) approach*, deserves some comments given its prominent role, for example, in the works of Esó and Szentes (2007, 2013), or Kakade et al (2011). Representing the type processes by means of independent shocks is always without loss of generality, and, as explained after Theorem 1, it provides a convenient way to establish primitive conditions under which the envelope formula is a necessary condition for incentive compatibility. However, the IS approach is not particularly useful for establishing (necessary and) sufficient conditions for implementability in Markov environments, because the transformation to independent shocks doesn’t in general preserve the Markovness of the environment. Hence, after the transformation, it is not sufficient to consider one-stage deviations from strongly truthful strategies (see the supplementary appendix for a counterexample). Accordingly, our analysis of implementability in Markov environments in Section 3.3 makes no reference to the IS approach.

Esó and Szentes (2007, 2013) have emphasized the fact that whenever the relaxed problem (or first-order) approach is valid, the cost to the principal of implementing a given allocation rule is the same as in a hypothetical environment where she can observe the agents’ future independent shocks (this is an immediate implication of the proof of Theorem 1, see footnote 33). Thus, in this sense, the agents do not receive rents on their orthogonal future information, suggesting an appealing intuition for profits and information rents based on the IS approach. It is worth noting, however, that this result holds—and the intuition is correct—only when the first-order approach is valid.<sup>39</sup>

The most important direction for future work pertains to the generality of our results on optimal dynamic mechanisms. In particular, our results were restricted to settings where the first-order approach yields an implementable allocation rule. To what extent this affects qualitative findings about the properties of optimal mechanisms is an open question. For some progress in this direction, see Garrett and Pavan (2013), who work directly with the integral monotonicity condition to show that, in the context of managerial compensation, the key properties of optimal contracts extend to environments where the first-order approach is invalid.

## A Proofs

**Proof of Theorem 1.** We start by establishing ICFOC<sub>*i*,0</sub> for all  $i = 1, \dots, n$ . Let the type processes be generated by the state representation  $\langle \mathcal{E}_i, G_i, z_i \rangle_{i=1}^n$  and consider a fictitious environment in which, in each period  $t \geq 1$ , each agent  $i = 1, \dots, n$  observes the shock  $\varepsilon_{it}$  and computes

---

<sup>39</sup>For example, the irrelevance result of Esó and Szentes (2013) does not hold if the solution to our relaxed problem is not implementable.

$\theta_{it} = Z_{i,(0),t}(\theta_{i0}, x_i^{t-1}, \varepsilon_i^t)$ . Consider a direct revelation mechanism in the fictitious environment in which each agent  $i$  reports  $\theta_{i0}$  in period 0 and  $\varepsilon_{it}$  in each period  $t \geq 1$ , and which implements the decision rule  $\hat{\chi}_t(\theta_0, \varepsilon^t) = \chi_t(Z_{(0)}^t(\theta_0, \hat{\chi}^{t-1}(\theta_0, \varepsilon^{t-1}), \varepsilon^t))$  and payment rule  $\hat{\psi}_t(\theta_0, \varepsilon^t) = \psi_t(Z_{(0)}^t(\theta_0, \hat{\chi}^{t-1}(\theta_0, \varepsilon^{t-1}), \varepsilon^t))$  in each period  $t$  (defined recursively on  $t$  with  $Z_{(0),t} \equiv (Z_{i,(0),t})_{i=1}^n$ ,  $Z_{(0)}^t = (Z_{(0),s})_{s=0}^t$ , and  $Z_{(0)} = (Z_{(0),s})_{s=0}^\infty$ ).

Suppose that all agents other than  $i$  report truthfully in all periods. Agent  $i$ 's payoff when the other agents' initial signals are  $\theta_{-i,0}$ , agent  $i$ 's true period-0 signal is  $\theta_{i0}$ , his period-0 report is  $\hat{\theta}_{i0}$ , and all future shocks  $\varepsilon$  are reported truthfully is given by

$$\hat{U}_i(\hat{\theta}_{i0}, \theta_{i0}, \theta_{-i,0}, \varepsilon) \equiv U_i(Z(\theta_{i0}, \theta_{-i,0}, \hat{\chi}(\hat{\theta}_{i0}, \theta_{-i,0}, \varepsilon), \varepsilon), \hat{\chi}(\hat{\theta}_{i0}, \theta_{-i,0}, \varepsilon)) + \sum_{t=0}^{\infty} \delta^t \hat{\psi}_{it}(\hat{\theta}_{i0}, \theta_{-i,0}, \varepsilon^t).$$

Since  $\langle \chi, \psi \rangle$  is PBIC in the original environment, truthful reporting by each agent at all truthful histories is a PBE of the mechanism  $\langle \hat{\chi}, \hat{\psi} \rangle$  in the fictitious environment. This implies that agent  $i$  cannot improve his expected payoff by misreporting his period-0 type and then reporting the subsequent shocks truthfully. That is, for any  $\theta_{i0} \in \Theta_{i0}$ ,

$$V_i^{(\chi, \psi)}(\theta_{i0}) = \sup_{\hat{\theta}_{i0} \in \Theta_{i0}} W(\hat{\theta}_{i0}, \theta_{i0}) = W(\theta_{i0}, \theta_{i0}), \text{ where } W(\hat{\theta}_{i0}, \theta_{i0}) \equiv \mathbb{E} \left[ \hat{U}_i(\hat{\theta}_{i0}, \theta_{i0}, \tilde{\theta}_{-i,0}, \tilde{\varepsilon}) \right].$$

The following lemma shows that the objective function  $W$  in the above maximization problem is well-behaved in the parameter  $\theta_{i0}$ :

**Lemma A.1** *Suppose that the environment is regular. Then, for all  $i = 1, \dots, n$  and  $\hat{\theta}_{i0} \in \Theta_{i0}$ ,  $W_i(\hat{\theta}_{i0}, \cdot)$  is equi-Lipschitz continuous and differentiable, with the derivative at  $\theta_{i0} = \hat{\theta}_{i0}$  given by*

$$\frac{\partial W_i(\hat{\theta}_{i0}, \hat{\theta}_{i0})}{\partial \theta_{i0}} = \mathbb{E} \left[ \sum_{t=0}^{\infty} \frac{\partial U_i \left( Z_{(0)} \left( \hat{\theta}_{i0}, \tilde{\theta}_{-i,0}, \hat{\chi} \left( \hat{\theta}_{i0}, \tilde{\theta}_{-i,0}, \tilde{\varepsilon} \right), \tilde{\varepsilon} \right), \hat{\chi} \left( \hat{\theta}_{i0}, \tilde{\theta}_{-i,0}, \tilde{\varepsilon} \right) \right)}{\frac{\partial Z_{i,(0),t} \left( \hat{\theta}_{i0}, \hat{\chi}_i^{t-1} \left( \hat{\theta}_{i0}, \tilde{\theta}_{-i,0}, \tilde{\varepsilon}^{t-1} \right), \tilde{\varepsilon}_i^t \right)}{\partial \theta_{i0}}} \right].$$

**Proof of Lemma A.1.** Let us focus on those  $\varepsilon$  for which  $\partial Z_{i,(0),t}(\theta_{i0}, \hat{\chi}_i^{t-1}(\hat{\theta}_{i0}, \theta_{-i,0}, \varepsilon_i^{t-1}), \varepsilon_i^t) / \partial \theta_{i0} < C_{it}(\varepsilon_i)$  for all  $i, t, \theta_0$ , with  $\|C_{i,(0)}(\varepsilon_i)\| < \infty$ , and  $\|Z_{i,(0)}(\theta_{i0}, \hat{\chi}_i(\theta_{i0}, \theta_{-i,0}, \varepsilon), \varepsilon_i)\| < \infty$ , which under Conditions F-BE and F-BIR occurs with probability 1, and temporarily drop arguments  $\varepsilon, \theta_{-i,0}, \hat{\theta}_{i0}, x = \hat{\chi}(\hat{\theta}_{i0}, \theta_{-i,0}, \varepsilon)$ , and subscripts  $i, (0)$  to simplify notation.

The classical chain rule (using Conditions U-D and F-BIR) yields that, for any given  $T$ ,

$$\Delta_T(\theta_0, h) \equiv \frac{1}{h} U(Z^T(\theta_0 + h), Z^{>T}(\theta_0)) - \frac{1}{h} U(Z(\theta_0)) - \sum_{t=0}^T \frac{\partial U(Z(\theta_0))}{\partial \theta_t} Z'_t(\theta_0) \rightarrow 0 \text{ as } h \rightarrow 0. \quad (17)$$

Note that

$$\frac{1}{h} U(Z^T(\theta_0 + h), Z^{>T}(\theta_0)) \rightarrow \frac{1}{h} U(Z(\theta_0 + h)) \text{ as } T \rightarrow \infty$$

uniformly in  $h$  since, using U-ELC, the difference is uniformly bounded by

$$A \frac{1}{h} \sum_{t=T+1}^{\infty} \delta^t |Z_t(\theta_0 + h) - Z_t(\theta_0)| \leq A \sum_{t=T+1}^{\infty} \delta^t C_t,$$

and the right-hand side converges to zero as  $T \rightarrow \infty$  since  $\|C\| < \infty$ .

Also, the series in (17) converges uniformly in  $h$  by the Weierstrass M-test, since, using Conditions U-ELC and F-BIR,

$$\sum_{t=0}^T \left| \frac{\partial U(Z(\theta_0))}{\partial \theta_t} \right| |Z'_t(\theta_0)| \leq \sum_{t=0}^T A \delta^t C_t \rightarrow A \|C\| \text{ as } T \rightarrow \infty$$

Hence, we have

$$\Delta_T(\theta_0, h) \rightarrow \frac{1}{h} [\hat{U}(\theta_0 + h) - \hat{U}(\theta_0)] - \sum_{t=0}^{\infty} \frac{\partial U(Z(\theta_0))}{\partial \theta_t} Z'_t(\theta_0) \text{ as } T \rightarrow \infty$$

uniformly in  $h$ . By uniform convergence we interchange the order of limits and use (17) to get

$$\begin{aligned} \lim_{h \rightarrow 0} \left[ \frac{1}{h} [\hat{U}(\theta_0 + h) - \hat{U}(\theta_0)] - \sum_{t=0}^{\infty} \frac{\partial U(Z(\theta_0))}{\partial \theta_t} Z'_t(\theta_0) \right] &= \lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \Delta_T(\theta_0, h) \\ &= \lim_{T \rightarrow \infty} \lim_{h \rightarrow 0} \Delta_T(\theta_0, h) = 0. \end{aligned}$$

This yields (putting back all the missing arguments)

$$\frac{\partial \hat{U}_i(\hat{\theta}_{i0}, \theta_{i0}, \varepsilon)}{\partial \theta_{i0}} = \sum_{t=0}^{\infty} \frac{\partial U_i(Z_{(0)}(\theta_0, \varepsilon, \hat{\chi}(\hat{\theta}_{i0}, \theta_{-i,0}, \varepsilon)), \hat{\chi}(\hat{\theta}_{i0}, \theta_{-i,0}, \varepsilon))}{\partial \theta_{it}} \frac{\partial Z_{i,(0),t}(\theta_{i0}, \varepsilon_i^t, \hat{\chi}_i^{t-1}(\hat{\theta}_{i0}, \theta_{-i,0}, \varepsilon^{t-1}))}{\partial \theta_{i0}}.$$

Next, note that, being a composition of Lipschitz continuous functions,  $\hat{U}_i(\hat{\theta}_{i0}, \cdot, \theta_{-i,0}, \varepsilon)$  is equi-Lipschitz continuous in  $\theta_{i0}$  with constant  $A \|C_{(0),i}(\varepsilon)\|$ . Since, by F-BIR,  $\mathbb{E}[\|C_{i,(0)}(\tilde{\varepsilon})\|] < \infty$ , by the Dominated Convergence Theorem we can write

$$\begin{aligned} \frac{\partial W_i(\hat{\theta}_{i0}, \theta_{i0})}{\partial \theta_{i0}} &= \lim_{h \rightarrow 0} \mathbb{E} \left[ \frac{\hat{U}_i(\hat{\theta}_{i0}, \theta_{i0} + h, \tilde{\theta}_{-i,0}, \tilde{\varepsilon}) - \hat{U}_i(\hat{\theta}_{i0}, \theta_{i0}, \tilde{\theta}_{-i,0}, \tilde{\varepsilon})}{h} \right] \\ &= \mathbb{E} \lim_{h \rightarrow 0} \left[ \frac{\hat{U}_i(\hat{\theta}_{i0}, \theta_{i0} + h, \tilde{\theta}_{-i,0}, \tilde{\varepsilon}) - \hat{U}_i(\hat{\theta}_{i0}, \theta_{i0}, \tilde{\theta}_{-i,0}, \tilde{\varepsilon})}{h} \right] = \mathbb{E} \left[ \frac{\partial \hat{U}_i(\hat{\theta}_{i0}, \theta_{i0}, \tilde{\theta}_{-i,0}, \tilde{\varepsilon})}{\partial \theta_{i0}} \right]. \end{aligned}$$

Hence, for any  $\hat{\theta}_{i0}$ ,  $W_i(\hat{\theta}_{i0}, \cdot)$  is differentiable and equi-Lipschitz in  $\theta_{i0}$  with Lipschitz constant  $A \mathbb{E}[\|C_{i,(0)}(\tilde{\varepsilon})\|]$ , and with derivative at  $\theta_{i0} = \hat{\theta}_{i0}$  given by the formula in the lemma. ■

The equi-Lipschitz continuity of  $W_i(\hat{\theta}_{i0}, \cdot)$  established in Lemma A.1 implies that the value function  $\sup_{\hat{\theta}_{i0} \in \Theta_{i0}} W(\hat{\theta}_{i0}, \theta_{i0})$ , which coincides with the equilibrium payoff  $V_i^{(X, \Psi)}(\theta_{i0})$ , is Lipschitz contin-

uous.<sup>40</sup> Furthermore, by Theorem 1 of Milgrom and Segal (2002), at any differentiability point of  $V_i^{(\chi, \Psi)}(\theta_{i0})$ , we have  $\frac{dV_i^{(\chi, \Psi)}(\theta_{i0})}{d\theta_{i0}} = \frac{\partial W_i(\theta_{i0}, \theta_{i0})}{\partial \theta_{i0}}$ . Using Lemma A.1, the Law of Iterated Expectations, and the definition of  $I_{i,(0),t}$  in (2) then yields  $\text{ICFOC}_{i,0}$ .  $\text{ICFOC}_{i,s}$  for  $s > 0$  then follows by the same argument, since agent  $i$ 's problem at a truthful period- $s$  history is identical to the period-0 problem except for the indexing by the history. ■

**Proof of Theorem 2.** For part (i), we show first that the flow transfers in (7) are well-defined by showing that the discounted sum (over  $t$ ) of each of the three terms in  $\psi_{it}$  has a finite expected net present value (NPV) under the measure  $\lambda_i[\chi, \Gamma]|\theta^{s-1}, \theta_{is}$  for all  $i = 1, \dots, n$ ,  $s \geq 0$ , and  $(\theta^{s-1}, \theta_{is}) \in \Theta^{s-1} \times \Theta_{is}$  (which implies that the series  $\sum_{t=0}^{\infty} \delta^t \psi_{it}(\theta^t)$  converges with probability 1 under  $\lambda_i[\chi, \Gamma]|\theta^{s-1}, \theta_{is}$ ). For the first term, using U-ELC and F-BIR, we have

$$\left| D_{it}^{\chi, \Gamma}(\theta^{t-1}, \theta_{it}) \right| \leq \mathbb{E}^{\lambda_i[\chi, \Gamma]|\theta^{t-1}, \theta_{it}} \left[ \sum_{\tau=t}^{\infty} \left| I_{i,(t),\tau}(\tilde{\theta}_i, \chi_i(\tilde{\theta})) \right| A_i \delta^\tau \right] \leq \delta^t A_i B_i, \quad (19)$$

where  $A_i > 0$  is the constant of equi-Lipschitz continuity of  $U_i$ , and where  $B_i > 0$  is the bound on the impulse responses in Condition F-BIR. This means that

$$\left| \delta^{-t} Q_{it}^{\chi, \Gamma}(\theta^{t-1}, \theta_{it}) \right| \leq A_i B_i |\theta_{it} - \theta'_{it}| \leq A_i B_i (|\theta_{it}| + |\theta'_{it}|). \quad (20)$$

Hence, the expected NPV of the first term is finite by Condition F-BE and  $\|\theta'_i\| < \infty$ . For the second term, using (20) for  $t+1$  and the Law of Iterated Expectations, the expected NPV of its absolute value is bounded by

$$A_i B_i \left( \sum_{\tau=0}^{t-1} \delta^{\tau+1} \mathbb{E}^{\lambda_i[\chi, \Gamma]|\theta^{\tau-1}, \theta_{i\tau}} [\tilde{\theta}_{i\tau+1}] + \mathbb{E}^{\lambda_i[\chi, \Gamma]|\theta^{s-1}, \theta_{is}} \left[ \|\tilde{\theta}_i\| \right] + \|\theta'_i\| \right),$$

which is finite by Condition F-BE and  $\|\theta'_i\| < \infty$ . Finally, the expected NPV of the third term is finite by Conditions U-SPR and F-BE.

We then show that  $\text{ICFOC}_{i,s}$  holds for all  $i$  and  $s$ . Rewrite the time- $s$  equilibrium expected payoff given history  $(\theta^{s-1}, \theta_{is})$  using Fubini's Theorem and the Law of Iterated Expectations as follows:

$$\begin{aligned} V_{is}^{(\chi, \Psi), \Gamma}(\theta^{s-1}, \theta_{is}) &= \lim_{T \rightarrow \infty} \sum_{t=0}^T \delta^t \mathbb{E}^{\lambda_i[\chi, \Gamma]|\theta^{s-1}, \theta_{is}} \left[ u_{it}(\tilde{\theta}^t, \chi^t(\tilde{\theta}^t)) + \psi_{it}(\tilde{\theta}^t) \right] \\ &= \sum_{t=0}^{s-1} \delta^t \left( \mathbb{E}^{\Gamma_i(\theta_i^{s-1}, \chi_i^{s-1}(\theta^{s-1}))} \left[ u_{it}(\theta_i^t, \tilde{\theta}_{-i}^t, \chi^t(\theta_i^t, \tilde{\theta}_{-i}^t)) \right] + \psi_{it}(\theta^t) \right) \\ &\quad + Q_{is}^{\chi, \Gamma}(\theta^{s-1}, \theta_{is}) - \lim_{T \rightarrow \infty} \mathbb{E}^{\lambda_i[\chi, \Gamma]|\theta^{s-1}, \theta_{is}} \left[ Q_{i,T+1}^{\chi, \Gamma}(\tilde{\theta}^T, \tilde{\theta}_{i,T+1}) \right]. \end{aligned}$$

---

<sup>40</sup>Since for each  $\theta_{i0}, \theta'_{i0}$ ,  $\left| V_i^{(\chi, \Psi)}(\theta'_{i0}) - V_i^{(\chi, \Psi)}(\theta_{i0}) \right| \leq \sup_{\hat{\theta}_{i0} \in \Theta_{i0}} \left| W_i(\hat{\theta}_{i0}, \theta'_{i0}) - W_i(\hat{\theta}_{i0}, \theta_{i0}) \right| \leq M |\theta'_{i0} - \theta_{i0}|$ , where  $M > 0$  is the constant of equi-Lipschitz continuity of  $W$ . This argument is similar to the first part of Milgrom and Segal's (2002) Theorem 2.

(The expectations of the other terms for  $t \geq s$  cancel out by the Law of Iterated Expectations). The second line is independent of  $\theta_{is}$ , and the limit on the last line equals zero by (20), Condition F-BE, and  $\|\theta'_i\| < \infty$ . By (6) and (19), the remaining term  $Q_{is}^{\chi, \Gamma}(\theta^{s-1}, \theta_{is})$  is Lipschitz continuous in  $\theta_{is}$ , and its derivative equals  $D_{is}^{\chi, \Gamma}(\theta^{s-1}, \theta_{is})$  a.e., which is the right-hand side of (1).

For part (ii), we start by considering a single-agent environment and then extend the result to multiple agents under the no-leakage condition.

Consider the single-agent case, where beliefs are vacuous, and omit the agent index to simplify notation. For any PBIC choice rules  $\langle \chi, \psi \rangle$  and  $\langle \chi, \bar{\psi} \rangle$  with the same allocation rule  $\chi$ , for all  $s \geq 0$  and  $\theta^s \in \Theta^s$ , ICFOC<sub>s</sub> and the Law of Iterated Expectations imply

$$\begin{aligned} V_s^{\langle \chi, \psi \rangle}(\theta^s) - V_{s-1}^{\langle \chi, \psi \rangle}(\theta^{s-1}) &= V_s^{\langle \chi, \psi \rangle}(\theta^s) - \mathbb{E}^{F_s}(\theta^{s-1}, \chi^{s-1}(\theta^{s-1})) \left[ V_s^{\langle \chi, \psi \rangle}(\theta^{s-1}, \tilde{\theta}_s) \right] \\ &= \mathbb{E}^{F_s}(\theta^{s-1}, \chi^{s-1}(\theta^{s-1})) \left[ \int_{\tilde{\theta}_s}^{\theta^s} D_s^{\chi}(\theta^{s-1}, q) dq \right] \\ &= V_s^{\langle \chi, \bar{\psi} \rangle}(\theta^s) - V_{s-1}^{\langle \chi, \bar{\psi} \rangle}(\theta^{s-1}). \end{aligned}$$

Substituting the definitions of expected payoffs and rearranging terms yields

$$\mathbb{E}^{\lambda[\chi]|\theta^s} \left[ \sum_{t=0}^{\infty} \delta^t \psi_t(\tilde{\theta}) \right] - \mathbb{E}^{\lambda[\chi]|\theta^s} \left[ \sum_{t=0}^{\infty} \delta^t \bar{\psi}_t(\tilde{\theta}) \right] = \mathbb{E}^{\lambda[\chi]|\theta^{s-1}} \left[ \sum_{t=0}^{\infty} \delta^t \psi_t(\tilde{\theta}) \right] - \mathbb{E}^{\lambda[\chi]|\theta^{s-1}} \left[ \sum_{t=0}^{\infty} \delta^t \bar{\psi}_t(\tilde{\theta}) \right].$$

By induction, we then have, for all  $T \geq 1$  and  $\theta^T \in \Theta^T$ ,

$$\mathbb{E}^{\lambda[\chi]|\theta^T} \left[ \sum_{t=0}^{\infty} \delta^t \psi_t(\tilde{\theta}) \right] - \mathbb{E}^{\lambda[\chi]|\theta^T} \left[ \sum_{t=0}^{\infty} \delta^t \bar{\psi}_t(\tilde{\theta}) \right] = \mathbb{E}^{\lambda[\chi]} \left[ \sum_{t=0}^{\infty} \delta^t \psi_t(\tilde{\theta}) \right] - \mathbb{E}^{\lambda[\chi]} \left[ \sum_{t=0}^{\infty} \delta^t \bar{\psi}_t(\tilde{\theta}) \right] \equiv K. \quad (21)$$

Since payoff from truth-telling in a PBIC mechanism is well-defined, we have the following lemma.

**Lemma A.2** *Suppose  $\psi$  is the transfer rule in a PBIC mechanism. Then for  $\lambda[\chi]$ -almost all  $\theta$ ,  $\mathbb{E}^{\lambda[\chi]|\theta^T} \left[ \sum_{t=0}^{\infty} \delta^t \psi_t(\tilde{\theta}) \right] \rightarrow \sum_{t=0}^{\infty} \delta^t \psi_t(\theta)$  as  $T \rightarrow \infty$ .*

**Proof of Lemma A.2.** By the Law of Iterated Expectations,

$$\begin{aligned} \mathbb{E}_{\tilde{\theta}}^{\lambda[\chi]} \left[ \left| \mathbb{E}_{\tilde{\theta}}^{\lambda[\chi]|\tilde{\theta}^T} \left[ \sum_{t=0}^{\infty} \delta^t \psi_t(\tilde{\theta}) \right] - \sum_{t=0}^{\infty} \delta^t \psi_t(\tilde{\theta}) \right| \right] &= \mathbb{E}_{\tilde{\theta}}^{\lambda[\chi]} \left[ \left| \mathbb{E}_{\tilde{\theta}}^{\lambda[\chi]|\tilde{\theta}^T} \left[ \sum_{t=T+1}^{\infty} \delta^t \psi_t(\tilde{\theta}^t) \right] - \sum_{t=T+1}^{\infty} \delta^t \psi_t(\tilde{\theta}^t) \right| \right] \\ &\leq 2 \mathbb{E}_{\tilde{\theta}}^{\lambda[\chi]} \left[ \sum_{t=T+1}^{\infty} \delta^t \left| \psi_t(\tilde{\theta}^t) \right| \right]. \end{aligned}$$

By PBIC,  $\mathbb{E}^{\lambda[\chi]}[|\psi(\tilde{\theta})|] < \infty$ , and hence the term on the second line goes to zero as  $T \rightarrow \infty$ . ■

By Lemma A.2, we can take the limit  $T \rightarrow \infty$  in (21) to get

$$\sum_{t=0}^{\infty} \delta^t \psi_t(\theta) - \sum_{t=0}^{\infty} \delta^t \bar{\psi}_t(\theta) = K \text{ for } \lambda[\chi]\text{-almost all } \theta.$$

In order to extend the result to multiple agents under the no-leakage condition, observe that if  $\langle \chi, \psi \rangle$  and  $\langle \chi, \bar{\psi} \rangle$  are PBIC, then they remain PBIC also in the “blind” setting where agent  $i$  does not observe his allocation  $x_i$ . (Hiding the allocation  $x_i$  from agent  $i$  simply amounts to pooling some of his incentive constraints.) Furthermore, if the allocation rule  $\chi$  leaks no information to agent  $i$  so that observing the true type  $\theta_i$  does not reveal any information about  $\theta_{-i}$ , then we can interpret the “blind” setting as a single-agent setting in which agent  $i$ 's allocation in period  $t$  is simply his report  $\hat{\theta}_{it}$ , and his utility is  $\hat{U}_i(\theta_i, \hat{\theta}_i) = \mathbb{E}^{\lambda_i[\chi]|\hat{\theta}_i} \left[ U_i(\theta_i, \tilde{\theta}_{-i}, \chi(\hat{\theta}_i, \tilde{\theta}_{-i})) \right]$ , where  $\lambda_i[\chi]|\hat{\theta}_i$  denotes the probability measure over the other agents' types when agent  $i$ 's reports are fixed at  $\hat{\theta}_i$ . (Intuitively, the other agents' types can be viewed as being realized only after agent  $i$  has finished reporting, and  $\hat{U}_i$  is the expectation taken over such realizations.) Applying to this setting the result established above for the single-agent case, we see that agent  $i$ 's expected payment  $\mathbb{E}^{\lambda_i[\chi]|\hat{\theta}_i} \left[ \sum_{t=0}^{\infty} \delta^t \psi_{it}(\theta_i, \tilde{\theta}_{-i}) \right]$  is pinned down, up to a constant, by the allocation rule  $\chi$  with probability 1. ■

**Proof of Theorem 3.** Given a choice rule  $\langle \chi, \psi \rangle$  and belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$ , for all  $i = 1, \dots, n$ ,  $t \geq 0$ , and  $(\theta_i^t, (\theta_i^{t-1}, \hat{\theta}_{it}), \theta_{-i}^{t-1}) \in \Theta_i^t \times \Theta_i^t \times \Theta_{-i}^{t-1}$ , let

$$\Phi_i(\theta_{it}, \hat{\theta}_{it}) \equiv V_{it}^{\langle \chi \circ \hat{\theta}_{it}, \psi \circ \hat{\theta}_{it} \rangle, \Gamma}(\theta^{t-1}, \theta_{it}).$$

That is,  $\Phi_i(\theta_{it}, \hat{\theta}_{it})$  denotes agent  $i$ 's expected payoff from reporting  $\hat{\theta}_{it}$  at the period- $t$  history  $(\theta_i^t, \theta_i^{t-1}, \chi_i^{t-1}(\theta_i^{t-1}, \theta_{-i}^{t-1}))$ , and then reverting to truthful reporting period  $t+1$  onwards. Because the environment is Markov,  $\Phi_i(\theta_{it}, \hat{\theta}_{it})$  depends on agent  $i$ 's past reports, but not on his past true types, and hence it gives agent  $i$ 's payoff from reporting  $\hat{\theta}_{it}$  at the history  $((\bar{\theta}_i^{t-1}, \theta_{it}), \theta_i^{t-1}, \chi_i^{t-1}(\theta_i^{t-1}, \theta_{-i}^{t-1}))$  for any past true types  $\bar{\theta}_i^{t-1} \in \Theta_i^{t-1}$ . We then let  $\bar{\Phi}_i(\theta_{it}) \equiv \Phi_i(\theta_{it}, \theta_{it})$  for all  $\theta_{it} \in \Theta_{it}$  denote the payoff from reporting truthfully in all periods  $s \geq t$ .

**Necessity:** Fix  $i = 1, \dots, n$  and  $t \geq 0$ . Suppose that the allocation rule  $\chi \in \mathcal{X}$  with belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$  can be implemented in an on-path truthful PBE. Then there exists a transfer rule  $\psi$  such that the choice rule  $\langle \chi, \psi \rangle$  with belief system  $\Gamma$  is PBIC, and thus satisfies ICFOC $_{i,t}$  by Theorem 1. This implies that  $\bar{\Phi}_i(\cdot)$  satisfies condition (b) in Lemma 1 with  $\bar{\Phi}'_i(\theta_{it}) = D_{it}^{\chi, \Gamma}(\theta^{t-1}, \theta_{it})$  for a.e.  $\theta_{it}$ . Thus it remains to establish condition (a).

**Lemma A.3** *Suppose the environment is regular and Markov. Fix  $i = 1, \dots, n$  and  $t \geq 0$ . If the choice rule  $\langle \chi, \psi \rangle$  with belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$  satisfies ICFOC $_{i,t+1}$ , then for all  $\hat{\theta}_{it} \in \Theta_{it}$ , the choice rule  $\langle \chi \circ \hat{\theta}_{it}, \psi \circ \hat{\theta}_{it} \rangle$  with belief system  $\Gamma$  satisfies ICFOC $_{i,t}$ .*

**Proof of Lemma A.3.** Note first that because the environment is Markov and  $\langle \chi, \psi \rangle$  with belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$  satisfies ICFOC $_{i,t+1}$ , the choice rule  $\langle \hat{\chi}, \hat{\psi} \rangle \equiv \langle \chi \circ \hat{\theta}_{it}, \psi \circ \hat{\theta}_{it} \rangle$  with belief system  $\Gamma$  clearly satisfies ICFOC $_{i,t+1}$  because agent  $i$ 's payoff does not depend on whether the previous period report  $\hat{\theta}_{it}$  has been truthful or not. In order to show that it also satisfies ICFOC $_{i,t}$ , we can use a state representation and the Law of Iterated Expectations to write agent  $i$ 's expected payoff from



truthtelling under choice rule  $\langle \hat{\chi}, \hat{\psi} \rangle$ , for all  $(\theta^{t-1}, \theta_{it})$ , as

$$V_{it}^{\langle \hat{\chi}, \hat{\Psi} \rangle, \Gamma}(\theta^{t-1}, \theta_{it}) = \mathbb{E}_{\tilde{\theta}_{-i}^t} \mathbb{E}_{\tilde{\varepsilon}_i^{t+1}} \left[ V_{i,t+1}^{\langle \hat{\chi}, \hat{\Psi} \rangle, \Gamma} \left( (\theta_{it}^t, \tilde{\theta}_{-i}^t), Z_{i,(t),t+1}(\theta_{it}, \hat{\chi}_i^t(\theta_{it}^t, \tilde{\theta}_{-i}^t), \tilde{\varepsilon}_i^{t+1}) \right) \right],$$

where  $\tilde{\theta}_{-i}^t$  is generated by drawing  $\theta_{-i}^{t-1}$  according to agent  $i$ 's belief  $\Gamma_i(\theta_{-i}^{t-1}, \chi_{-i}^{t-1}(\theta^{t-1}))$  and then drawing  $\tilde{\theta}_{-i,t}$  from  $\prod_{j \neq i} F_{jt}(\theta_j^{t-1}, \chi_j^{t-1}(\theta^{t-1}))$ . To differentiate this identity with respect to the true period- $t$  type  $\theta_{it}$ , note first that by the chain rule, we have

$$\begin{aligned} \frac{d}{d\theta_{it}} \left[ V_{i,t+1}^{\langle \hat{\chi}, \hat{\Psi} \rangle, \Gamma}(\theta^t, Z_{i,(t),t+1}(\theta_{it}, \hat{\chi}_i^t(\theta_{it}^t, \theta_{-i}^t), \varepsilon_{i,t+1})) \right] &= \mathbb{E}^{\lambda_i[\hat{\chi}, \Gamma]|\theta^t, \theta_{i,t+1}} \left[ \frac{\partial U_i(\tilde{\theta}, \hat{\chi}(\tilde{\theta}))}{\partial \theta_{it}} \right] \\ &+ D_{i,t+1}^{\hat{\chi}, \Gamma}(\theta^t, \theta_{i,t+1}) \frac{\partial Z_{i,(t),t+1}(\theta_{it}, \chi_i^t(\theta_{it}^t, \theta_{-i}^t), \varepsilon_{i,t+1})}{\partial \theta_{it}}. \end{aligned}$$

To see this, note that the first term follows because the environment is Markov and  $\langle \hat{\chi}, \hat{\psi} \rangle$  does not depend on  $\theta_{it}$  so that

$$\frac{\partial V_{i,t+1}^{\langle \hat{\chi}, \hat{\Psi} \rangle, \Gamma}(\theta^t, \theta_{i,t+1})}{\partial \theta_{it}} = \mathbb{E}^{\lambda_i[\hat{\chi}, \Gamma]|\theta^t, \theta_{i,t+1}} \left[ \frac{\partial U_i(\tilde{\theta}, \hat{\chi}(\tilde{\theta}))}{\partial \theta_{it}} \right].$$

The second term follows because, by ICFOC $_{i,t+1}$ ,  $\partial V_{i,t+1}^{\langle \hat{\chi}, \hat{\Psi} \rangle, \Gamma}(\theta^t, \theta_{i,t+1}) / \partial \theta_{i,t+1} = D_{i,t+1}^{\hat{\chi}, \Gamma}(\theta^t, \theta_{i,t+1})$ . Furthermore, by U-ELC, ICFOC $_{i,t+1}$ , and F-BIR all the derivatives above are bounded. Thus, by the Dominated Convergence Theorem, we can pass the derivative through the expectation to get

$$\begin{aligned} &\frac{dV_{it}^{\langle \hat{\chi}, \hat{\Psi} \rangle, \Gamma}(\theta^{t-1}, \theta_{it})}{d\theta_{it}} \\ &= \mathbb{E}^{\lambda_i[\hat{\chi}, \Gamma]|\theta^{t-1}, \theta_{it}} \left[ \frac{\partial U_i(\tilde{\theta}, \hat{\chi}(\tilde{\theta}))}{\partial \theta_{it}} + I_{i,(t),t+1}(\tilde{\theta}_i^{t+1}, \hat{\chi}_i^t(\theta_{it}^t, \theta_{-i}^t)) \sum_{\tau=t+1}^{\infty} I_{i,(t+1),\tau}(\tilde{\theta}_i^\tau, \hat{\chi}_i^{\tau-1}(\tilde{\theta})) \frac{\partial U_i(\tilde{\theta}, \hat{\chi}(\tilde{\theta}))}{\partial \theta_{i\tau}} \right] \\ &= \mathbb{E}^{\lambda_i[\hat{\chi}, \Gamma]|\theta^{t-1}, \theta_{it}} \left[ \sum_{\tau=t}^{\infty} I_{i,(t),\tau}(\tilde{\theta}_i^\tau, \hat{\chi}_i^{\tau-1}(\tilde{\theta})) \frac{\partial U_i(\tilde{\theta}, \hat{\chi}(\tilde{\theta}))}{\partial \theta_{i\tau}} \right], \end{aligned}$$

where we have first used (2) to express the expectation in terms of impulse responses, and then the fact that Markovness implies  $I_{i,(t),t+1}(\theta_i^{t+1}, x_i^t) I_{i,(t+1),\tau}(\theta_i^\tau, x_i^{\tau-1}) = I_{i,(t),\tau}(\theta_i^\tau, x_i^{\tau-1})$ . Therefore, the choice rule  $\langle \hat{\chi}, \hat{\psi} \rangle$  with belief system  $\Gamma$  satisfies ICFOC $_{i,t}$ . ■

Since  $\langle \chi, \psi \rangle$  with belief system  $\Gamma$  satisfies ICFOC by Theorem 1, Lemma A.3 implies that for all  $\hat{\theta}_{it} \in \Theta_{it}$ ,  $\langle \chi \circ \hat{\theta}_{it}, \psi \circ \hat{\theta}_{it} \rangle$  with belief system  $\Gamma$  satisfies ICFOC $_{i,t}$ . Therefore, for any fixed  $\hat{\theta}_{it}$ ,  $\Phi_i(\theta_{it}, \hat{\theta}_{it})$  is Lipschitz continuous in  $\theta_{it}$  with derivative given by  $D_{it}^{\chi \circ \hat{\theta}_{it}, \Gamma}(\theta^{t-1}, \theta_{it})$  for a.e.  $\theta_{it}$ . Hence, also condition (a) of Lemma 1 is satisfied. Since  $i$  and  $t$  were arbitrary, we conclude that the integral monotonicity condition (8) is a necessary condition for on-path truthful PBE implementability.

**Sufficiency:** Suppose the allocation rule  $\chi \in \mathcal{X}$  with belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$  satisfies integral monotonicity. Define the transfer rule  $\psi$  by (7). By Theorem 2, the choice rule  $\langle \chi, \psi \rangle$  with belief

system  $\Gamma$  satisfies ICFOC. Thus the above arguments show that, for all  $i = 1, \dots, n$ ,  $t \geq 0$ , and any period- $t$  history of agent  $i$ , the functions  $\{\Phi_i(\cdot, \hat{\theta}_{it})\}_{\hat{\theta}_{it} \in \Theta_{it}}$  and  $\bar{\Phi}_i(\cdot)$  satisfy conditions (a) and (b) of Lemma A.3. This implies that a one-step deviation from the strong truthtelling strategy is not profitable for agent  $i$  at *any* history in any period. The following version of the one-stage deviation principle then rules out multi-step deviations:

**Lemma A.4** *Suppose the environment is regular and Markov. Fix an allocation rule  $\chi \in \mathcal{X}$  with belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$ , and define the transfer rule  $\psi$  by (7). If a one-stage deviation from strong truthtelling is not profitable at any information set, then arbitrary deviations from strong truthtelling are not profitable at any information set.*

The proof of this lemma consists of showing that despite payoffs being not a priori continuous at infinity, the bounds implied by U-SPR and part (ii) of the definition of Markov environments guarantee that under the transfers defined by (7), continuation utility is well-behaved. We relegate the argument to the supplementary material.

We conclude that integral monotonicity is a sufficient condition for the allocation rule  $\chi \in \mathcal{X}$  with belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$  to be implementable in a strongly truthful PBE. ■

**Proof of Proposition 1.** Case (i): We construct a nondecreasing solution  $\chi_s(\theta)$  sequentially for  $s = 0, 1, \dots$ . Suppose we have a solution  $\chi$  in which  $\chi^{s-1}(\theta)$  is nondecreasing. Consider the problem of choosing the optimal continuation allocation rule in period  $s$  given type history  $\theta^s$  and allocation history  $\chi^{s-1}(\theta^{s-1})$ . Using the state representation  $\langle \mathcal{E}_i, G_i, z_i \rangle_{i=1}^n$  from period  $s$  onward, we can write the continuation rule for  $t \geq s$  as a collection of functions  $\hat{\chi}_t(\varepsilon)$  of the shocks  $\varepsilon$ .

First, note that, because  $X$  is a sublattice,  $\prod_{t \geq s} X_t$  is a lattice. This means that the set of feasible shock-contingent plans  $\hat{\chi}$  is also a lattice under pointwise meet and join operations (i.e., for each  $\varepsilon$ ). Next, note that, under the assumptions in the proposition, each agent  $i$ 's virtual utility

$$U_i(Z_s(\theta^s, \varepsilon), \chi^{s-1}(\theta^{s-1}), x^{\geq s}) - \frac{1}{\eta_{i0}(\theta_{i0})} \sum_{t=0}^{\infty} \frac{\partial U_i(Z_s(\theta^s, \varepsilon), \chi^{s-1}(\theta^{s-1}), x^{\geq s})}{\partial \theta_{it}} I_{i,(0),t}(Z_{i,(s)}^t(\theta^s, \varepsilon))$$

is supermodular in  $x^{\geq s}$  and has increasing differences in  $(\theta^s, x^{\geq s})$  (note that  $Z_s(\theta^s, \varepsilon)$  is nondecreasing in  $\theta^s$  by F-FOSD, and  $\chi^{s-1}(\theta^{s-1})$  is nondecreasing in  $\theta^{s-1}$  by construction). Therefore, summing over  $i$  and taking expectation over  $\varepsilon$ , we obtain that the expected virtual surplus starting with history  $\theta^s$  is supermodular in the continuation plan  $\hat{\chi}$  and has increasing differences in  $(\theta^s, \hat{\chi})$ . Topkis's Theorem then implies that the set of optimal continuation plans is nondecreasing in  $\theta^s$  in the strong set order. In particular, focus on the first component  $\chi_s \in X_s$  of such plans. By Theorem 2 of Kukushkin (2009), there exists a nondecreasing selection of optimal values,  $\hat{\chi}_s(\theta^s)$ . Therefore, the relaxed program has a solution in which  $\chi^s(\theta^s) = (\chi^{s-1}(\theta^{s-1}), \hat{\chi}_s(\theta^s))$  is nondecreasing in  $\theta^s$ .

Case (ii): In this case, the solution to the relaxed problem is a collection of independent rules  $\chi_t$ ,

$t \geq 0$ , with each  $\chi_t$  satisfying, for  $\lambda$ -almost every  $\theta^t$ ,

$$\chi_t(\theta^t) \in \arg \max_{x_t \in X_t} \left[ \sum_{i=0}^n u_{it}(\theta_t, x_t) - \sum_{i=1}^n \frac{1}{\eta_{i0}(\theta_{i0})} I_{i,(0),t}(\theta_i^t) \frac{\partial u_{it}(\theta_t, x_t)}{\partial \theta_{it}} \right]. \quad (22)$$

For any  $t$ , we can put  $\chi_t(\theta^t) = \bar{\chi}_t(\varphi_{1t}(\theta_1^t), \dots, \varphi_{nt}(\theta_n^t))$  for some  $\bar{\chi}_t: \mathbb{R}^{n \times m} \rightarrow X_t$ . Now fix  $i \geq 1$ , and for  $x_{it} \in X_{it}$ , let  $X_t(x_{it}) \equiv \{x'_t \in X_t : x'_{it} = x_{it}\}$ . Then (22) implies

$$\bar{\chi}_{it}(\varphi_t) \in \arg \max_{x_{it} \in X_{it}} [\bar{u}_{it}(\varphi_{it}, x_{it}) + g_{it}(\varphi_{-i,t}, x_{it})]$$

where  $\bar{u}_{it}(\varphi_{it}, x_{it})$  is the virtual utility of agent  $i$ , and

$$g_{it}(\varphi_{-i,t}, x_{it}) \equiv \max_{x'_t \in X_t(x_{it})} \left[ u_{0t}(x'_t) + \sum_{j \neq i} \bar{u}_{jt}(\varphi_{jt}, x'_{jt}) \right]$$

Since  $\bar{u}_{it}(\varphi_{it}, x_{it}) + g_{it}(\varphi_{-i,t}, x_{it})$  has strictly increasing differences in  $(\varphi_{it}, x_{it})$ , by the Monotone Selection Theorem of Milgrom and Shannon (1994),  $\bar{\chi}_{it}(\varphi_{it}, \varphi_{-i,t})$  must be nondecreasing in  $\varphi_{it}$ , and so  $\chi_{it}(\theta_i^t, \theta_{-i}^t)$  is nondecreasing in  $\theta_i^t$ . ■

**Proof of Proposition 2.** Fix a belief system  $\Gamma \in \mathbf{\Gamma}(\chi)$ . We show that the virtual index policy given by (16) satisfies average monotonicity: For all  $i = 1, \dots, n$ ,  $s \geq 0$ , and  $(\theta^{s-1}, \theta_{is}) \in \Theta^{s-1} \times \Theta_{is}$ ,

$$\mathbb{E}^{\lambda_i[\chi \circ \hat{\theta}_s, \Gamma] | \theta^{s-1}, \theta_{is}} \left[ \sum_{t=s}^{\infty} \delta^t (\chi \circ \hat{\theta}_{is})_{it}(\tilde{\theta}) \right]$$

is nondecreasing in  $\hat{\theta}_{is}$ . We show this for  $s = 0$ . The argument for  $s > 0$  is analogous but simpler since  $\theta_{is}$  does not affect the term  $\eta_{i0}^{-1}(\theta_{i0})$  in the definition of the virtual index (15) when  $s > 0$ .

We can think of the processes being generated as follows: First, draw a sequence of innovations  $\omega_i = (\omega_{ik})_{k=1}^{\infty}$  according to  $\prod_{k=1}^{\infty} R_i(\cdot | k)$  for each  $i$ , independently across  $i = 1, \dots, n$ , and draw initial types  $\theta_{i0}$  according to  $F_{i0}$  independently of the innovations  $\omega_i$  and across  $i$ . Letting  $K_t \equiv \sum_{\tau=1}^t x_{\tau}$ , bidder  $i$ 's type in period  $t$  can then be described as

$$\theta_{it} = \theta_{i0} + \sum_{k=1}^{K_t} \omega_{ik}.$$

Clearly this representation generates the same conditional distributions (and hence the same process) as the kernels defined in the main text.<sup>41</sup>

<sup>41</sup>The difference is that in this representation, the innovation to bidder  $i$ 's value if he wins the auction for the  $k$ th time in period  $t$  is given by the  $k$ th element of the sequence  $\omega_i$ , whereas in the representation in the main text it is given by (a function of) the  $t$ th element of the sequence  $\varepsilon_i$ . In terms of the latter, a higher current message increases discounted consumption only on average, whereas in the former it increases discounted consumption for each realization of  $\omega_i$ , since the bidder always experiences the same innnovation sequence irrespective of the timing of consumption.

Next, fix an arbitrary bidder  $i = 1, \dots, n$  and a state  $(\theta_0, \omega) \in \Theta_0 \times (\mathbb{R}^n)^\infty$ , and take a pair  $\theta'_{i0}, \theta''_{i0} \in \Theta_0$  with  $\theta''_{i0} > \theta'_{i0}$ . We show by induction on  $k$  that, for any  $k \in \mathbb{N}$ , the  $k$ th time that  $i$  wins the object if he initially reports  $\theta''_{i0}$  (and reports truthfully in each period  $t > 0$ ) comes weakly earlier than if he reports  $\theta'_{i0}$ . As the realization  $(\theta_0, \omega) \in \Theta_0 \times (\mathbb{R}^n)^\infty$  is arbitrary, this implies that the expected time to the  $k$ -th win is decreasing in the initial report, which in turn implies that the virtual policy  $\chi$  satisfies average monotonicity.

As a preliminary observation, note that the period- $t$  virtual index of bidder  $i$  is increasing in the (reported) period-0 type  $\theta_{i0}$  since the handicap  $\eta_{i0}^{-1}(\theta_{i0})$  is decreasing in  $\theta_{i0}$ , and (in case  $t = 0$ )  $\mathbb{E}^{\lambda[\bar{\chi}_i]|\theta_{i0}}[\theta_{i\tau}]$  is increasing in  $\theta_{i0}$  for all  $\tau \geq 0$ .

Base case: Suppose, towards a contradiction, that the first win given initial report  $\theta'_{i0}$  comes in period  $t'$  whereas it comes in period  $t'' > t'$  given report  $\theta''_{i0} > \theta'_{i0}$ . As the realization  $(\theta_0, \omega)$  is fixed, the virtual indices of bidders  $-i$  in period  $t'$  are the same in both cases. But  $\gamma_{it'}((\theta''_{i0}, \theta_{i0}, \dots, \theta_{i0}), 0) > \gamma_{it'}((\theta'_{i0}, \theta_{i0}, \dots, \theta_{i0}), 0)$ , implying that  $i$  must win in period  $t'$  also with initial report  $\theta''_{i0}$ , which contradicts  $t'' > t'$ .

Induction step: Suppose the claim is true for some  $k \geq 1$ . Suppose towards contradiction that the  $(k + 1)$ th win given report  $\theta'_{i0}$  comes in period  $t'$  whereas it comes in period  $t'' > t'$  given  $\theta''_{i0} > \theta'_{i0}$ . Then observe that (i) In both cases,  $i$  wins the auction  $k - 1$  times prior to period  $t'$ . Furthermore, since the realization  $(\theta_0, \omega)$  is fixed, this implies that (ii) bidder  $i$ 's current type  $\theta_{it}$  is the same in both cases, and (iii) the number of times each bidder  $j \neq i$  wins the object prior to period  $t'$  is the same in both cases, and hence the virtual indices of bidders  $-i$  in period  $t'$  are the same in both cases. By (i) and (ii)  $i$ 's virtual index in period  $t'$  is identical in both cases except for the initial report. That bidder  $i$ 's period- $t'$  index is increasing in the initial report, along with (iii) implies that  $i$  must then win in period  $t'$  also with initial report  $\theta''_{i0}$ , contradicting  $t'' > t'$ . Hence the claim is true for  $k + 1$ . ■

## References

- [1] Albanesi, S. and C. Sleet (2006), "Dynamic Optimal Taxation with Private Information," *Review of Economic Studies*, 73(1), 1–30.
- [2] Athey, S. and I. Segal (2013), "An Efficient Dynamic Mechanism," *Econometrica*, 81(6), 2463–2485.
- [3] Atkeson, A. and R. E. Jr Lucas (1992), "On efficient Distribution with Private Information," *Review of Economic Studies*, 59(3), 427–453.
- [4] Baron, D. and D. Besanko (1984), "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, 1, 447–470.

- [5] Battaglini, M. (2005), “Long-Term Contracting with Markovian Consumers,” *American Economic Review*, 95(3), 637–658.
- [6] Battaglini, M. (2007), “Optimality and Renegotiation in Dynamic Contracting,” *Games and Economic Behavior*, 60(2), 213–246.
- [7] Battaglini, M. and R. Lamba (2012), “Optimal Dynamic Contracting,” Mimeo, Princeton University.
- [8] Besanko, D. (1985) “Multiperiod Contracts Between Principal and Agent with Adverse Selection,” *Economics Letters*, 17, 33–37.
- [9] Bergemann, D. and M. Said (2011), “Dynamic Auctions,” *Wiley Encyclopedia of Operations Research and Management Science*.
- [10] Bergemann, D. and J. Välimäki (2008), “Bandit Problems,” *New Palgrave Dictionary of Economics*, Durlauf and Blume eds.
- [11] Bergemann, D. and J. Välimäki (2010), “The Dynamic Pivot Mechanism,” *Econometrica*, 78(2), 771–789.
- [12] Berger, A., R Müller, and S. H. Naeemi (2010), “Path-Monotonicity and Truthful Implementation,” Mimeo, Maastricht University.
- [13] Billingsley, P. (1995), *Probability and Measure*, 3rd ed., Wiley.
- [14] Board, S. (2007), “Selling Options,” *Journal of Economic Theory* 136, 324–340.
- [15] Cremer, J. and R. P. McLean (1988), “Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions,” *Econometrica*, 56(6), 1247–1257.
- [16] Courty, P. and H. Li (2000), “Sequential Screening,” *Review of Economic Studies*, 67, 697–717.
- [17] Dudley, R. (2002), *Real Analysis and Probability*, Cambridge University Press.
- [18] Carbajal, J.C. and J. Ely (2013), “Mechanism Design without Revenue Equivalence,” *Journal of Economic Theory*, 148(1), 104–133.
- [19] Eső, P. and B. Szentes (2007), “Optimal Information Disclosure in Auctions and the Handicap Auction,” *Review of Economic Studies*, 74(3), 705–731.
- [20] Eső, P. and B. Szentes (2013), “Dynamic Contracting with Adverse Selection: An Irrelevance Result,” Mimeo, LSE and Oxford University.

- [21] Farhi, E. and I. Werning (2013), “Insurance and Taxation over the Life Cycle,” *Review of Economic Studies*, 80(2), 596–635.
- [22] Fernandes, A. and C. Phelan (2000), “A Recursive Formulation for Repeated Agency with History Dependence,” *Journal of Economic Theory*, 91(2), 223–247.
- [23] Fudenberg, D. and J. Tirole (1991), *Game Theory*, MIT Press.
- [24] Garcia, D. (2005), “Monotonicity in Direct Revelation Mechanisms,” *Economic Letters*, 88(1), 21–26.
- [25] Garrett, D. (2011), “Durable-good Sales with Changing Values and Unobservable Arrival Dates,” Mimeo, Northwestern.
- [26] Garrett, D. and A. Pavan (2013), “Dynamic Managerial Compensation: On the Optimality of Seniority-Based Schemes,” Mimeo, Northwestern University.
- [27] Gershkov, A. and B. Moldovanu (2009), “Dynamic Revenue Maximization with Heterogenous Objects: A mechanism Design Approach,” *American Economic Journal: Microeconomics*, 1(2), 168–198.
- [28] Golosov, M., M. Troschkin, and A. Tsyvinski (2011), “Optimal Dynamic Taxes,” Mimeo, Yale University.
- [29] Green, E. J. (1987), “Lending and the Smoothing of Uninsurable Income,” in Prescott E. and Wallace N., eds, *Contractual Agreements for Intertemporal Trade*, University of Minnesota Press.
- [30] Hendel, I. and A. Lizzeri (2003), “The Role Of Commitment In Dynamic Contracts: Evidence From Life Insurance,” *Quarterly Journal of Economics*, 118(1), 299–327.
- [31] Kakade, S., Lobel, I., and H. Nazerzadeh (2011), “Optimal Dynamic Mechanism Design and the Virtual Pivot Mechanism,” Mimeo Wharton School, University of Pennsylvania.
- [32] Kapicka, M. (2013), “Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-Order Approach,” *Review of Economic Studies*, 80(3), 1027–1054.
- [33] Kocherlakota, N. R. (2005), “Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation,” *Econometrica*, 73(5), 1587–1621.
- [34] Kukushkin, N.S. (2009), “On the Existence of Monotone Selections,” Mimeo.
- [35] Laffont, J.-J. and J. Tirole (1988), “The Dynamics of Incentive Contracts,” *Econometrica*, 56(5), 1153–1175.

- [36] Matthews, S. and J. Moore (1987), “Monopoly provision of quality and warranties: An exploration in the theory of multidimensional screening.” *Econometrica*, 55(2), 441–467.
- [37] Milgrom, P. and I. Segal (2002), “Envelope Theorems for Arbitrary Choice Sets,” *Econometrica*, 70(2), 583–601.
- [38] Mirrlees, J.A. (1971), “An Exploration in the Theory of Optimum Income Taxation,” *Review of Economic Studies*, 38(114), 175–208.
- [39] Pollard, D. (2002), *A User’s Guide to Measure Theoretic Probability*, Cambridge University Press.
- [40] Rahman, D. (2010), “Dynamic Implementation,” Mimeo, University of Minnesota.
- [41] Riordan, M., and D. Sappington (1987), “Awarding Monopoly Franchises,” *American Economic Review*, 77(3), 375–387.
- [42] Rochet, J. C. (1987), “A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context,” *Journal of Mathematical Economics*, 16, 191–200.
- [43] Rochet, J-C. and L. Stole (2003), “The Economics of Multidimensional Screening,” in M. Dewatripont, L.P. Hansen, and S. Turnovsky, eds, *Advances in Economics and Econometrics: Theory and Applications - Eight World Congress of Econometric Society*, 150–197. Cambridge University Press.
- [44] Skreta, V. (2006), “Sequentially Optimal Mechanisms,” *Review of Economic Studies*, 73(4), 1085–1111.
- [45] Skrzypacz, A. and J. Toikka (2013), “Mechanisms for Repeated Bargaining,” Mimeo, MIT and Stanford.
- [46] Strausz, R. (2006), “Deterministic versus Stochastic Mechanisms in Principal-Agent Models,” *Journal of Economic Theory*, 127(1), 306–314.
- [47] Strulovici, B. (2011), “Renegotiation-proof Contracts with Moral Hazard and Persistent Private Information,” Mimeo, Northwestern University.
- [48] Whittle, P. (1982), *Optimization Over Time, vol. 1*, Wiley, Chichester.
- [49] Williams, N. (2011), “Persistent Private Information,” *Econometrica*, 79(4), 1233–1275.