Long-term Contracting in a Changing World*

Alessandro Pavan†

First version: May 2007. This version: December 2007
("still" incomplete)

Abstract

I study the properties of optimal long-term contracts in an environment in which the agent's type evolves stochastically over time. The model stylizes a buyer-seller relationship but the results apply quite naturally to many contractual situations including regulation and optimal income-taxation. I first show, through a simple discrete example, that distortions need not vanish over time and need not be monotonic in the shock to the buyer's valuation. These results are in contrast to those obtained in the literature that assumes a Markov process with a binary state space—e.g. Battaglini, 2005. I then show that the study of the dynamics of the optimal mechanism can be significantly simplified by assuming the shocks are independent over time. When the sets of possible types in any two adjacent periods satisfy a certain overlapping condition (which is always satisfied with a continuum of types) and some additional regularity conditions hold, then the optimal mechanism is the same irrespective of whether the shocks are the buyer's private information or are observed also by the seller. These conditions are satisfied, for example, in the case of an AR(1) process, a Brownian motion, but also when shocks have a multiplicative effect as it is often the case in financial applications. Furthermore, the distortions in the optimal quantities are independent of the distributions of the shocks and, when the buyer's payoff is additively separable, they are also independent of whether the shocks are transitory or permanent. Finally, I show that assuming the shocks are independent not only does it greatly simplify the analysis, it is actually without loss of generality.

Keywords: asymmetric information, stochastic process, dynamic mechanism design, long-term contracting

JEL classification numbers: D82, C73, L1

*I thank Marco Battaglini, Igal Hendel, Phil Reny, Bill Rogerson, Marciano Siniscalchi, Balazs Szentes, Michael Whinston, and seminar participants at Universitad Carlos III de Madrid, the University of Chicago, and Northwestern University for comments and useful suggestions. I also thank the Department of Economics at the University of Chicago for hospitality during the fall of 2007 when part of this work was completed. Keziah Cook and Sorin Maruster provided excellent research assistance.

†Department of Economics, Northwestern University. Email: alepavan@northwestern.edu

1 Introduction

Long-term contracting plays an important role in a variety of situations including trade, employment, regulation, taxation, and financial contracting. Most long-term relationships take place in a "changing world" that is in an environment that evolves (stochastically) over time: the value a buyer attaches to a good or a service, the production cost incurred by a seller, the productivity of a worker, the return to a financial project, the parties' outside options, are all likely to change over time in response to shocks to the environment. These changes are often anticipated at the moment of contracting, albeit not necessarily jointly observed by the parties. By implication, optimal long-term contracts must be flexible to accommodate such changes and be designed in a way that provides the parties with incentives to share the information they gradually learn over time.

Understanding the properties of optimal long-term contracts is important both from a positive and a normative viewpoint. Unfortunately, the characterization of optimal long-term contracts for stochastic environments can be tedious which possibly explains why the literature on optimal dynamic contracts is relatively thin and most of the existing results established only for special cases.

The purpose of this paper is to suggest a convenient way of describing the evolution of the environment which facilitates the characterization of optimal long-term contracts. I then use such a characterization to address the following questions: What are the dynamics of the optimal decisions, such as the supply of quantity/quality over time? When do distortions vanish in the long-run? Under what conditions does it matter whether the "shocks" (i.e. the changes to the environment) are privately observed instead of being jointly observed by the parties? When is the nature of the shocks (i.e. whether they are transitory or permanent) irrelevant for the dynamics of the optimal decisions?

The model I consider stylizes a dynamic buyer-seller relationship. However, the techniques and the results apply more generally to many other environments, including the applications mentioned above.

In the first part of the paper, I use a simple example to argue that the results obtained in the literature that assumes a Markov process with a binary state space—e.g. Battaglini, 2005 and Battaglini and Coate, 2007—need not extend to more general stochastic processes, such as a Brownian motion, or an AR(1) process. In particular, I show that distortions need not vanish in the long-run, can be larger in the future than in the present and need not be monotonic in the shocks to the buyer's valuation.

In the second part of the paper, I then consider environments in which the shocks are *independent* over time (but where types are correlated!). Assuming independent shocks greatly simplifies the analysis.

I first consider the case where the buyer's valuation (his type) follows a continuous process and then the case of finitely many types. In both cases, the approach I follow to characterize the optimal mechanism is the following. First, I derive necessary conditions for incentive-compatibility that are the analog in a dynamic setting of the familiar envelope conditions for static environments. Next, I define a *relaxed program* that consists of maximizing the seller's intertemporal expected payoff under the sole constraints that the mechanism satisfy the necessary conditions for incentive-compatibility. Thanks to the assumption of *independent shocks*, the solution to the relaxed program is extremely simple and illuminative. Letting θ_1 denote the buyer's type in period one, $\xi^t = (\xi_2, ..., \xi_t)$ the vector of shocks experienced by the buyer up to period t, and $v_t(\theta_1, \xi^t)$ the buyer's period-t valuation, the dynamics of distortions in the optimal mechanism are governed entirely by the dynamics of $\partial v_t(\theta_1, \xi^t)/\partial\theta_1$, i.e. by the sensitivity of future valuations to the buyer's initial type. In particular, the dynamics of distortions are completely independent of the distributions of the shocks. For example, when the agent's valuation follows a *random walk* (or, in continuos time, a Brownian motion), then $\theta_t = v_t(\theta_1, \xi^t) = \theta_1 + \xi_2 + \cdots + \xi_t$, in which case distortions are constant over time and depend only on the hazard rate of the distribution of the buyer's first-period valuation. In contrast, in the case of an AR(1) process, $\theta_t = v_t(\theta_1, \xi^t) = a_t\theta_{t-1} + \xi_t$ in which case distortions are governed by the dynamics of $\prod a_t$. If $a_t = a$ for all t with $a \in (0, 1)$, then distortions eventually vanish over time, as predicted by Battaglini's model. More generally, however, distortions need not be monotonic neither in time nor in the magnitude of the shocks and need not vanish in the long-run.

I also show that, when the optimal mechanism solves the relaxed program, then it has the following properties: (i) the quantity schedules (as well as the players' payoffs) coincide with the ones the seller would offer if the shocks were jointly observed by both parties; (ii) the seller may find it optimal to exclude a buyer for a few periods and then serve him again once his virtual valuation has sufficiently improved; (iii) high-valuation buyers may receive smaller quantities than low-valuation ones, as a function of how their valuations evolved over time; (iv) it is never optimal for the seller to transfer the ownership of the production technology to the buyer. Once again all these properties are in contrast to those that one obtains in a model with a binary type space.

Because the aforementioned properties refer to the solution to the relaxed program, it is important to understand under what conditions the latter coincides with the optimal mechanism. In other words, when are the necessary conditions also sufficient? In static environments the answer is know to rest upon the combination of two properties: the fact that the allocations are monotonic in the agent's type and the fact that the agent's preferences satisfy the *single-crossing property.* Unfortunately, in a dynamic setting, these properties alone do not guarantee that the solution to the relaxed program is indeed incentive-compatible.

The sufficient conditions I identify are the analog of those in Eso and Szentes (2007) adapted to the multi-period-multi-decision setting considered here. These conditions are based on the following two properties. That the process is *Markov* and that the quantity schedules satisfy a *strong monotonicity* condition. This condition requires that, holding constant the evolution of the buyer's type up to period $t-1$ and the value of the buyer's valuation from period $t+1$ to period $s \geq t+1$, the quantity supplied in period s be monotonic in the shock (equivalently, the buyer's type) in period t.

Following essentially the same reverse-engineering as in Eso and Szentes (2007), I then show how one can back up the primitive conditions for the underlying stochastic process that guarantee that the solution to the relaxed program satisfies the strong monotonicity property. The advantage of describing the evolution of the state (i.e. the buyer's type) through a sequence of independent innovations (as opposed to a sequence of conditional distributions, as it is standard in the literature—e.g. Baron and Besanko, 1984, Courty and Li, 2000) stems also from the possibility of identifying such sufficient conditions.

Next, I consider stochastic processes with finitely-many types. In this case, the aforementioned properties (Markov + strong monotonicity) do not suffice for the optimal mechanism to coincide with the solution to the relaxed program. If, however, in addition to these properties, a certain *overlapping condition* holds which requires that the set of possible types in any two adjacent periods overlap enough, then the same results established for the case with a continuum of types obtain: the optimal mechanism coincides with the one the seller would offer if the buyer could lie only once, as it is the case with observable shocks.

The natural question to ask at this point is how restrictive is the assumption of independent shocks? I show that any process in which the distribution of θ_t given $(\theta_1, ..., \theta_{t-1})$ is continuous and strictly increasing in θ_t can be represented through a collection of real valued functions v_t and random variables ξ_t such that $\theta_t = v_t(\theta_1, \xi^t)$ with (θ_1, ξ^t) jointly independent. Assuming independent shocks is thus without loss of generality in the case of a continuous process. The characterization of the necessary conditions for incentive-compatibility and the corresponding solution to the relaxed program are thus general. On the other hand, the sufficient conditions described above also require the process to be Markov, the functions $v_t(\theta_1, \xi^t)$ to be increasing in each argument, and the quantity schedules that solve the relaxed program to be strongly monotonic. The monotonicities of v_t are always satisfied under first-order-stochastic-dominance. Strong monotonicity of the schedules is satisfied when, in addition to the aforementioned properties, the distribution of the agent's type in period one is log-concave and the conditional distributions $\mathcal{F}_t(\theta_t; \theta_{t-1})$ are such that $[\partial \mathcal{F}_t(\theta_t; \theta_{t-1})/\partial \theta_{t-1}] \,/\, [f_t(\theta_t; \theta_{t-1})]$ are increasing in both θ_t and θ_{t-1}, a property that is satisfied by many continuous processes. Assuming independent shocks is thus not only convenient, but actually less restrictive than one may think.

The rest of the paper is organized as follows. Section 2 contains the model. Section 3 illustrates, by means of a simple (finite) example, why the results obtained with binary types need not extend to more general stochastic processes. Section 4 contains all the results for independent shocks. Finally, Section 5 discusses in what sense the assumption of independent shocks is not restrictive.

2 The environment

Consider a buyer-seller relationship that evolves over $T \in \mathbb{N}$ periods. Both the buyer and the seller are risk-neutral and have preferences represented, respectively, by

$$U = \sum_{t=1}^{T} \delta^{t-1}[\theta_t q_t - p_t] \quad \text{and} \quad \Pi = \sum_{t=1}^{T} \delta^{t-1}[p_t - C_t(q_t)]$$

where $q_t \in \mathbb{R}_+$ denotes the quantity received by the buyer in period t, θ_t the buyer's period-t valuation, p_t the *total* price paid to the seller in period t, $C_t(q_t)$ the cost to produce and supply q_t, and $\delta > 0$ the common discount factor. The function $C_t : \mathbb{R}_+ \longrightarrow \mathbb{R}$ is strictly increasing and convex with $C_t'(0) = 0$ and $\lim_{q\to\infty} C_t'(q) = +\infty$, for all t.

The evolution of the buyer's valuation is conveniently described through a collection of functions $v_t : \mathbb{R}^t \longrightarrow \mathbb{R}$ such that, for any $t > 1$,

$$\theta_t = v_t(\theta_1, \xi^t)$$

where θ_1 denotes the buyer's valuation in period one while $\xi^t \equiv (\xi_2, ..., \xi_t)$ denotes the vector of "shocks" experienced by the buyer in the subsequent periods. Each function v_t is equi-Lipschitz continuous, strictly increasing, and twice continuously differentiable, in each argument. The joint distribution of $\left(\theta_1, \xi^T\right)$ is described by the c.d.f. Ψ with support $\Theta_1 \times \Xi_2 \times \cdots \times \Xi_T$. With an abuse of terminology, I will hereafter refer to $\theta_t \in \Theta_t$ as the buyer's "type" in period t, with

$$\Theta_t \equiv \{\theta \in \mathbb{R} : \theta = v_t(\theta_1, \xi^t),\ (\theta_1, \xi^t) \in \Theta_1 \times \Xi_2 \times \cdots \times \Xi_t\}.$$

To simplify the description of the distortions in the optimal contracts, and without any serious loss of generality, hereafter I assume that $C_t(q) = q^2/2$ for all t, in which case the first-best schedules are given by

$$q_t^{FB}(\theta_1, \xi^t) = v_t(\theta_1, \xi^t) \ \forall t \text{ and } \forall \left(\theta_1, \xi^t\right)$$

The sequence of events is the following.

- At $t = 0$, the buyer privately learns θ_1.
- At $t = 1$, the seller offers a mechanism $\varphi = (\mathcal{M}, \phi)$. The latter consists of a collection of mappings

$$\phi_t : \mathcal{M}_1 \times \cdots \times \mathcal{M}_t \to \mathbb{R}_+ \times \mathbb{R}$$

  that specify a price-quantity pair for each possible profile of messages $m^t \equiv (m_1, ..., m_t) \in \mathcal{M}_1 \times \cdots \times \mathcal{M}_t$, with$\mathcal{M} \equiv (\mathcal{M})_{t=1}^T$ and $\phi \equiv (\phi_t)_{t=1}^T$. A mechanism is thus equivalent to a menu of long-term contracts with memory.

  If the buyer refuses to participate in φ, the game ends and both players obtain a payoff equal to zero. If the buyer accepts to participate in φ, he chooses a message $m_1 \in \mathcal{M}_1$, receives a quantity $q_1(m_1)$, pays a transfer $p_1(m_1)$, and the game moves to period 2.
- At $t = 2$, the buyer privately observes the shock ξ_2. He then chooses whether to continue to participate in φ or "walk away" from the relationship. If he walks away, the game ends and both players' continuation payoffs are zero. If he stays, he sends a new message $m_2 \in \mathcal{M}_2$, receives a new quantity $q_2(m_1, m_2)$, pays a transfer $p_2(m_1, m_2)$, and the game moves to period 3.
- $\cdots$
- At $t = T + 1$ the game is over.

Remark. The game described above corresponds to an environment with one-side commitment: the seller can perfectly commit to her long-term mechanism, but the buyer can walk away from the relationship at any point in time.

2.1 Incentive-compatibility

Because the seller can fully commit, the Revelation Principle[1] applies and, by implication, the profit-maximizing mechanism can be characterized by restricting attention to direct mechanisms in which the agent truthfully reveals his type θ_t at any point in time. Alternatively, one can restrict attention to mechanisms in which the agent is asked to report his initial type θ_1 in period one and the innovation ξ_t at any subsequent period. Clearly, by virtue of the Revelation Principle, the two approaches are equivalent. For reasons that will become clear from the subsequent exposition, I find the latter approach more convenient. In what follows, I thus consider direct mechanisms

$$\phi_t : \Theta_1 \times \Xi_2 \times \cdots \times \Xi_t \to \mathbb{R}_+ \times \mathbb{R}$$

[1] See, among others, Gibbard (1977), Green and Laffont (1977) and Myerson (1979).

in which the buyer reports at each period the innovation ξ_t instead of his type θ_t.[2] A direct mechanism is *incentive-compatible* if, conditional on having reported truthfully in the past, the buyer has the incentives to report truthfully in the present.[3]

Given the mechanism $\varphi = (\mathcal{M}, \phi)$, let $\tilde{U}_t\left(\theta_1, \xi^t; \hat{\theta}_1, \hat{\xi}^{t-1}\right)$ denote the continuation payoff for a buyer whose initial type was θ_1, who experienced the shocks ξ^t and who reported $\left(\hat{\theta}_1, \hat{\xi}^{t-1}\right)$ in the past. Assuming the function $\tilde{U}_t$ solves the Bellman equation, then

$$\begin{aligned}\tilde{U}_t\left(\theta_1, \xi^t; \hat{\theta}_1, \hat{\xi}^{t-1}\right) &= \max_{\hat{\xi}_t}\Big\{v_t(\theta_1, \xi^t)q_t(\hat{\theta}_1, \hat{\xi}^{t-1}, \hat{\xi}_t) - p_t(\hat{\theta}_1, \hat{\xi}^{t-1}, \hat{\xi}_t) \\ &\quad + \delta\mathbb{E}_{\xi_{t+1}}\left[\tilde{U}_{t+1}\left(\theta_1, \xi^t, \xi_{t+1}; \hat{\theta}_1, \hat{\xi}^{t-1}, \hat{\xi}_t\right) \mid \theta_1, \xi^t\right]\Big\}\end{aligned}$$

Now let $U_t\left(\theta_1, \xi^t\right)$ denote the buyer's expected *continuation* payoff from period t onward when he reported truthfully in the past and plans to report truthfully in all subsequent periods (i.e. his equilibrium path continuation payoff). Incentive compatibility then requires that, for any t and any $\left(\theta_1, \xi^t\right)$,

$$U_t\left(\theta_1, \xi^t\right) = \tilde{U}_t\left(\theta_1, \xi^t; \theta_1, \xi^{t-1}\right).$$

That is, a buyer who reported truthfully up to period $t-1$ and who experiences a shock ξ_t in period t, (weakly) prefers to report truthfully from period t onward rather than lying in period t and then choosing optimally what to report at any subsequent date.

The two-period case. As an illustration, suppose $T = 2$. Because there is no risk of confusion, I then drop the subscript for the shock in period two.

First, consider the buyer's incentives at $t = 2$. Let

$$U_2\left(\theta_1, \xi\right) \equiv v_2(\theta_1, \xi)q_2(\theta_1, \xi) - p_2(\theta_1, \xi) \tag{1}$$

denote the buyer's continuation payoff at $t = 2$ when he reported θ_1 truthfully at $t = 1$ and he reports ξ truthfully at $t = 2$. Then let

$$\tilde{U}_2\left(\theta_1, \xi; \hat{\theta}_1\right) \equiv \max_{\hat{\xi}}\left\{v_2(\theta_1, \xi)q_2(\hat{\theta}_1, \hat{\xi}) - p_2(\hat{\theta}_1, \hat{\xi})\right\}$$

[2] Note that $\left(\theta_1, \xi^t\right)$ conveniently summarizes not only the buyer's current valuation, θ_t,but the entire sequence of past valuations, $\theta^t \equiv (\theta_1, ..., \theta_t)$.

[3] One could also consider a stronger notion of incentive compatibility according to which the buyer has the incentives to report truthfully in each period, *regardless* of whether he has reported truthfully in the past. Because of the Revelation Principle, the two notions lead to the same optimal allocations. In fact, imposing incentive-compatibility also off-equilibrium is just a way of describing the best action for the buyer, given *any* possible history. As I show below, such a complete description is however unnecessary to establish the results.

denote the maximal continuation payoff that the buyer can guarantee himself in the (sub)game that starts at $t = 2$ when his true type in period one was θ_1, he reported $\hat{\theta}_1$, and then experienced a shock ξ in period two. Incentive-compatibility at $t = 2$ requires that

$$U_2(\theta_1, \xi) = \tilde{U}_2(\theta_1, \xi; \theta_1) \ \forall\ (\theta_1, \xi). \tag{2}$$

That is, conditional on having reported θ_1 truthfully in period one, the buyer finds it optimal to report ξ truthfully in period two.

Next, consider the buyer's incentives at $t = 1$. Let

$$U(\theta_1) \equiv \theta_1 q_1(\theta_1) - p_1(\theta_1) + \delta \mathbb{E}_\xi \{U_2(\theta_1, \xi) \mid \theta_1\} \tag{3}$$

denote the buyer's expected payoff when his type is θ_1 and he plans to report truthfully in both periods. Then let

$$\tilde{U}(\theta_1) \equiv \max_{\hat{\theta}_1} \left\{ \theta_1 q_1(\hat{\theta}_1) - p_1(\hat{\theta}_1) + \delta \mathbb{E}_\xi \{\tilde{U}_2\left(\theta_1, \xi; \hat{\theta}_1\right) \mid \theta_1\} \right\}$$

denote the maximal payoff that type θ_1 can obtain by choosing optimally his reports in each period. Incentive-compatibility at $t = 1$ then requires that

$$\theta_1 \in \arg\max_{\hat{\theta}_1} \left\{ \theta_1 q_1(\hat{\theta}_1) - p_1(\hat{\theta}_1) + \delta \mathbb{E}_\xi \{\tilde{U}_2\left(\theta_1, \xi; \hat{\theta}_1\right) \mid \theta_1\} \right\} \tag{4}$$

A mechanism is thus incentive-compatible if and only if (4) holds for any θ_1 and (2) holds for any (θ_1, ξ); equivalently, if and only if

$$U(\theta_1) = \tilde{U}(\theta_1) \ \forall \theta_1. \tag{5}$$

3 A simple (finite) example

To illustrate the trade-offs that determine the structure of the optimal mechanism in the simplest possible way, consider the following environment in which $T = 2$, $\Theta_1 \equiv \{\bar{\theta}, \underline{\theta}\}$, $\underline{\theta} > 0$, $\Delta\theta \equiv \bar{\theta} - \underline{\theta} > 0$, and $\Xi \equiv \{\xi_l, \xi_m, \xi_h\}$, with $\xi_h > \xi_m > \xi_l$. The probability the buyer is a high type (equivalently, the proportion of high types in the the cross section of the population) is $\Pr(\theta_1 = \bar{\theta}) = v$. Conditional on θ_1, the probability of a high shock is $\Pr(\xi_h | \theta_1) = x(\theta_1)$, the probability of an intermediate shock is $\Pr(\xi_m | \theta_1) = \alpha(\theta_1)$ and the probability of a low shock is $\Pr(\xi_l | \theta_1) = 1 - \alpha(\theta_1) - x(\theta_1)$, with $\bar{x} \equiv x(\bar{\theta})$, $\bar{\alpha} = \alpha(\bar{\theta})$, $\underline{\alpha} \equiv \alpha(\underline{\theta})$ and $\underline{x} \equiv x(\underline{\theta})$. For future reference, also let $\Delta\bar{\xi} \equiv \xi_h - \xi_m$,

$\Delta\underline{\xi} \equiv \xi_m - \xi_l$ and $\Delta\xi \equiv \xi_h - \xi_l$. Finally, without loss of generality, assume that

$$\theta_2 = v_2(\theta_1, \xi_2) = \theta_1 + \xi_2.$$

This simple environment suffices to illustrate the logic of the results; as I show below, it also nests Battaglini (2005) as a special case thus permitting us to illustrate what drives the results in his paper and why they need not extend to more general processes.

In this environment, incentive-compatibility at $t = 2$ requires that

$$U_2(\theta_1, \xi_h) - U_2(\theta_1, \xi_m) \geq \Delta\bar{\xi} q_2(\theta_1, \xi_m) \tag{6}$$

$$U_2(\theta_1, \xi_h) - U_2(\theta_1, \xi_l) \geq \Delta\xi q_2(\theta_1, \xi_l) \tag{7}$$

$$U_2(\theta_1, \xi_m) - U_2(\theta_1, \xi_l) \geq \Delta\underline{\xi} q_2(\theta_1, \xi_l) \tag{8}$$

$$U_2(\theta_1, \xi_h) - U_2(\theta_1, \xi_m) \leq \Delta\bar{\xi} q_2(\theta_1, \xi_h) \tag{9}$$

$$U_2(\theta_1, \xi_m) - U_2(\theta_1, \xi_l) \leq \Delta\underline{\xi} q_2(\theta_1, \xi_m) \tag{10}$$

$$U_2(\theta_1, \xi_h) - U_2(\theta_1, \xi_l) \leq \Delta\xi q_2(\theta_1, \xi_h) \tag{11}$$

both for $\theta_1 = \bar{\theta}$ and for $\theta_1 = \underline{\theta}$. From (6)-(11) one can immediately see that a necessary condition for incentive-compatibility at $t = 2$ is the familiar monotonicity condition according to which $q_2(\theta_1, \xi)$ is nondecreasing in ξ. The following lemma further simplify the characterization of incentive-compatibility.

Lemma 1 *Take any pair of types $\theta_2'' = v_2(\theta_1, \xi'')$ and $\theta_2' = v_2(\theta_1, \xi')$, with $\theta_2'' > \theta_2'$ and suppose that the period-2 contracts for these two types $\phi_2(\theta_1, \xi'') \equiv (q_2(\theta_1, \xi''), p_2(\theta_1, \xi''))$ and $\phi_2(\theta_1, \xi') \equiv (q_2(\theta_1, \xi'), p_2(\theta_1, \xi'))$, with $\phi_2(\theta_1, \xi'') \neq \phi_2(\theta_1, \xi')$, are designed in an incentive-compatible way, in the sense that*

$$\theta_2'' q_2(\theta_1, \xi'') - p_2(\theta_1, \xi'') \geq \theta_2'' q_2(\theta_1, \xi') - p_2(\theta_1, \xi') \tag{12}$$

$$\theta_2' q_2(\theta_1, \xi') - p_2(\theta_1, \xi') \geq \theta_2' q_2(\theta_1, \xi'') - p_2(\theta_1, \xi''). \tag{13}$$

Then,

(i) *Any type $\theta_2 > \theta_2''$ strictly prefers the contract $\phi_2(\theta_1, \xi'')$ to the contract $\phi_2(\theta_1, \xi')$.*

(ii) *Any type $\theta_2 < \theta_2'$ strictly prefers the contract $\phi_2(\theta_1, \xi')$ to the contract $\phi_2(\theta_1, \xi'')$.*

(iii) *Any type $\theta_2 \in (\theta_2', \theta_2'')$ strictly prefers the contract $\phi_2(\theta_1, \xi')$ to the contract $\phi_2(\theta_1, \xi'')$ if (12) binds and $\phi_2(\theta_1, \xi'')$ to $\phi_2(\theta_1, \xi')$ if (13) binds.*

Lemma 1 follows directly from the fact that the function $g(\theta, q) \equiv \theta \cdot q$ satisfies the *single crossing property* and thus has strictly increasing differences. An immediate implication of Lemma

1 is that, when the adjacent downstream local incentive-compatibility constraints (6) and (8) bind (or, alternatively, when the adjacent upstream constraints (9) and (10) bind) and $q_2(\theta_1, \xi)$ is monotonic in ξ, then all remaining constraints for $t = 2$ are satisfied. Lemma 1 also permits us to characterize the behavior of the buyer in period two when he misreported his type in period one.

Consider the payoff that $\bar{\theta}$ expects from lying at $t = 1$ and then reporting ξ truthfully at $t = 2$. The following is then a necessary condition for incentive-compatibility at $t = 1$.

$$\begin{aligned} U(\bar{\theta}) \geq\ & U(\underline{\theta}) + \Delta\theta q_1(\underline{\theta}) + \delta\{\bar{x}[U_2(\underline{\theta}, \xi_h) + \Delta\theta q_2(\underline{\theta}, \xi_h)] + \overline{\alpha}[U_2(\underline{\theta}, \xi_m) + \Delta\theta q_2(\underline{\theta}, \xi_m)] \\ & + (1 - \overline{\alpha} - \overline{x})[U_2(\underline{\theta}, \xi_l) + \Delta\theta q_2(\underline{\theta}, \xi_l)] \\ & - \underline{x}U_2(\underline{\theta}, \xi_h) - \underline{\alpha}U_2(\underline{\theta}, \xi_m) - (1 - \underline{\alpha} - \underline{x})U_2(\underline{\theta}, \xi_l)\} \end{aligned} \tag{14}$$

Condition (14) can be conveniently rewritten as

$$\begin{aligned} U(\bar{\theta}) \geq\ & U(\underline{\theta}) + \Delta\theta q_1(\underline{\theta}) + \delta\{(\overline{x} - \underline{x})[U_2(\underline{\theta}, \xi_h) - U_2(\underline{\theta}, \xi_m)] \\ & + (\bar{x} + \bar{\alpha} - \underline{\alpha} - \underline{x})[U_2(\underline{\theta}, \xi_m) - U_2(\underline{\theta}, \xi_l)] + \overline{x}\Delta\theta q_2(\underline{\theta}, \xi_h) \\ & + \overline{\alpha}\Delta\theta q_2(\underline{\theta}, \xi_m) + (1 - \overline{\alpha} - \overline{x})\Delta\theta q_2(\underline{\theta}, \xi_l)\} \end{aligned} \tag{15}$$

Now suppose that shocks are sufficiently "large" in the sense that $\Delta\theta \leq \min\{\Delta\bar{\xi}, \Delta\underline{\xi}\}$, so that $\bar{\theta} + \xi_l \leq \underline{\theta} + \xi_m$ and $\bar{\theta} + \xi_m \leq \underline{\theta} + \xi_h$ and that the distribution of θ_2 given $\bar{\theta}$ first-order-stochastic dominates the distribution of θ_2, given $\underline{\theta}$, i.e. $\overline{x} \geq \underline{x}$ and $\overline{x} + \overline{\alpha} \geq \underline{x} + \underline{\alpha}$. Neglecting the other incentive-compatibility constraints, it is then immediate that, to limit the rent of the high type, it is optimal for the seller to set $U(\underline{\theta}) = 0$ and to make the following downward adjacent incentive-compatibility constraints binding

$$\begin{aligned} U_2(\underline{\theta}, \xi_h) - U_2(\underline{\theta}, \xi_m) &\geq \Delta\bar{\xi} q_2(\underline{\theta}, \xi_m) \\ U_2(\underline{\theta}, \xi_m) - U_2(\underline{\theta}, \xi_l) &\geq \Delta\underline{\xi} q_2(\underline{\theta}, \xi_l) \end{aligned} \tag{16}$$

Note that, what determines the expected surplus of the high type are not the values of the continuation payoffs $U_2(\underline{\theta}, \xi)$ per se, but the "speed" at which the continuation payoffs for the low type changes with the shock ξ. Now suppose that, when searching for the optimal mechanism, the only relevant constraints are the participation constraint for the low type at $t = 1$, $U(\underline{\theta}) \geq 0$, the incentive-compatibility constraints (16) for the low type at $t = 2$ and the incentive-compatibility constraint (14) for the high type at $t = 1$ (any individually-rational and incentive-compatible mechanism must satisfy these constraints, but alone they do not necessarily guarantee that the buyer finds it optimal to participate and truthfully reveal his information in both periods). We then have the following result.

Lemma 2 *Assume*

$$\Delta\theta \leq \min\{\Delta\overline{\xi}, \Delta\underline{\xi}\},\ \bar{x} \geq \underline{x} \text{ and } \bar{x} + \overline{\alpha} \geq \underline{x} + \underline{\alpha}. \tag{17}$$

Any solution to the relax program

$$\mathcal{P}_r : \begin{cases} \max_{\phi} \mathbb{E}_{\theta_1} \left\{\theta_1 q_1(\theta) - C(q_1(\theta_1)) + \delta \mathbb{E}_{\xi}\left[(\theta_1 + \xi) q_2(\theta_1, \xi) - C(q_2(\theta_1, \xi)) \mid \theta_1\right] - U(\theta_1)\right\} \\ \text{subject to } U\left(\underline{\theta}\right) \geq 0,\ \text{(15) and (16)} \end{cases}$$

is such that all constraints in $\mathcal{P}_r$ bind and is characterized by the following quantities:

$$\begin{aligned} q_1(\overline{\theta}) &= q_1^{FB}(\overline{\theta}); \\ q_2(\overline{\theta}, \xi) &= q_2^{FB}(\overline{\theta}, \xi)\ \forall \xi \\ q_1(\underline{\theta}) &= \max\{q^{FB}(\underline{\theta}) - \tfrac{\nu}{1-\nu}\Delta\theta; 0\}; \\ q_2(\underline{\theta}, \xi_h) &= \max\{q^{FB}(\underline{\theta}, \xi_h) - \left(\tfrac{\nu}{1-\nu}\right)\left(\tfrac{\overline{x}}{\underline{x}}\right)\Delta\theta; 0\}; \\ q_2(\underline{\theta}, \xi_m) &= \max\{q^{FB}(\underline{\theta}, \xi_m) - \left(\tfrac{\nu}{1-\nu}\right)\left(\tfrac{(\overline{x}-\underline{x})\Delta\overline{\xi} + \overline{\alpha}\Delta\theta}{\underline{\alpha}}\right); 0\}; \\ q_2(\underline{\theta}, \xi_l) &= \max\{q^{FB}(\underline{\theta}, \xi_l) - \left(\tfrac{\nu}{1-\nu}\right)\left(\tfrac{(\overline{x}+\overline{\alpha}-\underline{x}-\underline{\alpha})\Delta\underline{\xi} + (1-\overline{\alpha}-\overline{x})\Delta\theta}{1-\underline{\alpha}-\underline{x}}\right); 0\}. \end{aligned}$$

The result in Lemma (2) is quite intriguing, for it suggests that, contrary to what indicated in the literature, distortions in the contract for the low type need not decrease over time. But when does the solution to the relaxed program actually coincide with the optimal mechanism, i.e. with the solution to the following unrelaxed program?

$$\mathcal{P} : \begin{cases} \max_{\phi} \mathbb{E}_{\theta_1} \left\{\theta_1 q_1(\theta) - C(q_1(\theta_1)) + \delta \mathbb{E}_{\xi}\left[(\theta_1 + \xi) q_2(\theta_1, \xi) - C(q_2(\theta_1, \xi)) \mid \theta_1\right] - U(\theta_1)\right\} \\ \text{subject to} \\ U(\theta_1) = \tilde{U}(\theta_1)\ \forall \theta_1 \qquad \text{(IC)} \\ U\left(\theta_1\right) \geq 0\ \forall \theta_1 \qquad \text{(IR-1)} \\ U_2\left(\theta_1, \xi\right) \geq 0\ \forall (\theta_1, \xi) \qquad \text{(IR-2)} \end{cases}$$

The answer is in the following proposition.

Proposition 1 *Assume Condition (17) holds. The solution to the relaxed program $\mathcal{P}_r$ coincides with the solution to the unrelaxed program $\mathcal{P}$ if and only if the quantities $q_2(\underline{\theta}, \xi)$ in Lemma (2) are nondecreasing in ξ. When this is the case, the optimal mechanism has the following properties:*

1. *There are no distortions in the contract for the high type;*

2. *All quantities in the contract for the low type are downward distorted;*

3. *Distortions in period two are (weakly) higher than those in period one;*

4. *Distortions need not be monotonic in the magnitude of the shock ξ. In particular, distortions can be higher after a favorable than an unfavorable shock to the buyer's type.*

It is useful to contrast the result in the previous proposition to the one in Battaglini (2005) for the case $\theta_t \in \{\bar{\theta}, \underline{\theta}\}$ for all t.

Proposition 2 (Battaglini) *Assume $\theta_t \in \{\bar{\theta}, \underline{\theta}\}$ for all t. The optimal contract has the following properties.*

(a) Generalized No Distortion at the Top: *As soon as the buyer experiences a positive shock which raises his valuation to $\bar{\theta}$, he receives first-best quantities thereafter.*

(b) Vanishing Distortion at the Bottom: *Distortions for the low type decrease over time and vanish in the long-run.*

Battaglini's model is nested in this example by letting $\Delta\overline{\xi} = \Delta\underline{\xi} = \Delta\theta$ and then letting $\bar{x} = 0 = 1 - \underline{\alpha} - \underline{x}$. The comparison between Propositions (1) and (2) uncovers interesting properties of optimal long-term contracts. A form of Generalized No Distortion at the Top seems robust: as soon as the buyer's valuation reaches the maximal possible value for period t, then the seller provides first-best output to the buyer thereafter. On the contrary, the Vanish Distortion at the Bottom does not appear robust; it is sensitive to the assumption that high types cannot experience shocks that further increase their valuations and likewise that low types cannot experience shocks that further reduce their valuations. Such a property thus holds when the buyer's valuation follows a Markov-1 process with a binary state space, but need not extend to more general Markov processes, as I further discuss in the next section.[4] Together these results offer an important message: What determines the dynamics of distortions is the familiar trade-off between efficiency and rent-extraction, *evaluated from period one's perspective*; not the fact that the agent's current type is just a noisy predictor of the agent's future types, as suggested in the literature. Actually, this can be seen directly from Battaglini's model by considering the following transition probabilities: $\Pr(\bar{\theta}|\bar{\theta}) = 1$ and $\Pr(\bar{\theta}|\underline{\theta}) = \alpha \in (0, .5)$. The distortions in the contract for the low type are then constant over time, despite the absence of perfect correlation. This result can also be seen from Proposition (1) by letting $\Delta\overline{\xi} = \Delta\underline{\xi} = \Delta\theta$. In this case, stochastic dominance simply requires that $\bar{x} + \bar{\alpha} > \underline{x}$. The

[4]The results that distortions need not decrease over time and need not be monotonic in the shocks extend to environments in which shocks are "small" as compared to the initial type θ_1, i.e. to settings in which condition (17) is violated.

Also note that, while in this paper I am assuming the seller can perfectly commit to her mechanism, the aforementioned results extend to settings in which the seller can offer long-term contracts but cannot commit not to renegotiate. Both results are available upon request.

solution to the optimal mechanism then coincides with the one in Lemma (2) substituting $\Delta\overline{\xi}$ and $\Delta\underline{\xi}$ with $\Delta\theta$. There are no distortions in the contract for the high type, whereas the quantities in the contract for the low type are given by

$$\begin{aligned} q_1(\underline{\theta}) &= \max\{q_1^{FB}(\underline{\theta}) - \tfrac{\nu}{1-\nu}\Delta\theta; 0\} \\ q_2(\underline{\theta}, \xi_h) &= \max\{q_2^{FB}(\underline{\theta}, \xi_h) - \left(\tfrac{\nu}{1-\nu}\right)\left(\tfrac{\overline{x}}{\underline{x}}\right)\Delta\theta; 0\} \\ q_2(\underline{\theta}, \xi_m) &= \max\{q_2^{FB}(\underline{\theta}, \xi_m) - \left(\tfrac{\nu}{1-\nu}\right)\left(\tfrac{\overline{x}+\overline{\alpha}-\underline{x}}{\underline{\alpha}}\right)\Delta\theta; 0\} \\ q_2(\underline{\theta}, \xi_l) &= \max\{q_2^{FB}(\underline{\theta}, \xi_l) - \left(\tfrac{\nu}{1-\nu}\right)\Delta\theta; 0\} \end{aligned} \tag{18}$$

The special case in which $\theta_t \in \{\bar{\theta}, \underline{\theta}\}$ for all t is then nested with $\bar{x} = 0 = 1 - \underline{\alpha} - \underline{x}$. In this case there is no distortion for $q_2(\underline{\theta}, \xi_h)$, whereas the quantity the seller supplies in period two to a low type whose valuation remains low is

$$q_2(\underline{\theta}, \xi_m) = \max\{q_2^{FB}(\underline{\theta}, \xi_m) - \left(\tfrac{\nu}{1-\nu}\right)\left(\tfrac{\overline{\alpha}+\underline{\alpha}-1}{\underline{\alpha}}\right)\Delta\theta; 0\}$$

When $\overline{\alpha} = 1$, i.e. when the valuation of the high type is constant over time, $q_2(\underline{\theta}, \xi_m) = q_1(\underline{\theta})$, irrespective of $\underline{\alpha}$, i.e. irrespective of the level of correlation between the buyer's valuations in the two periods.

Another case of special interest is when the buyer's types are correlated over time, but the shocks to the buyer's valuations are independent of θ, i.e. when $\bar{x} = \underline{x}$ and $\bar{\alpha} = \underline{\alpha}$. As one can immediately see from Lemma (2), in this case distortions in the contract for the low type are constant over time and do not depend on the shocks' distributions, i.e. on α and x. As I show in the next section, this property holds more generally whenever there is enough overlap in the support of the distribution of the agent's type over time—a property that is always satisfied with a continuum of types.

4 Independent shocks

From now on, consider the case in which ξ_t is independent of both θ_1 and ξ_s, for any $s \neq t$. A special case of interest (for many applications) is the case in which the buyer's valuation evolves in continuous time following a Brownian motion (possibly with drift), but in which trade occurs in discrete time. This case is nested in the model considered in this section. More generally, as I show in Section 5, all *continuous* Markov processes can be reconducted to the class considered in this section by appropriately specifying the distributions of the shocks and the v_t functions.

In what follows, I first consider the case where both θ_1 and ξ_t are drawn from a continuous distribution, and then turn back to the case of finitely-many types at the end of the section.

4.1 Continuum of types

Suppose that θ_1 is drawn from an absolutely continuous cumulative distribution function F with log-concave density f strictly positive over $\Theta_1 \equiv [\underline{\theta}, \bar{\theta}]$. Next, assume that, for any $t \geq 2$,

$$\theta_t = v_t(\theta_1, \xi^t)$$

with ξ_t drawn from an absolutely continuous cumulative distribution function G_t with support $\Xi_t \equiv [\underline{\xi}_t, \bar{\xi}_t]$. While for convenience I am restricting attention to environments with a compact state space, the results extend to environments in which the supports of F and G_t coincide with the entire real line as it is the case with Normally distributed shocks.

The functions v_t are assumed to be strictly increasing and twice continuously differentiable in each argument. A special case of interest is that of a (possibly non-stationary) AR(1) process:

$$\begin{aligned} \theta_t &= a_t\theta_{t-1} + \xi_t \\ &= \prod_{j=2}^{t} a_j\theta_1 + \prod_{j=3}^{t} a_j\xi_2 + \cdots + a_t\xi_{t-1} + \xi_t \end{aligned} \tag{19}$$

with $a_j \in \mathbb{R}_+$ for all $j \geq 2$. The case of independent types is then nested as $a_j = 0$, for all j, while the random-walk case is nested as $a_j = 1$.

To illustrate the logic behind the characterization results, assume for a moment that $T = 2$ (to save on notation, I then suppress the subscripts for v and ξ). The strategy I follow to characterize the optimal mechanism is the same as in the previous section. First, I identify some necessary conditions for incentive-compatibility; next, I maximize the seller's expected revenue subject to these conditions. Finally, I identify properties of the stochastic process that guarantee that these conditions are also sufficient.

First, consider incentive-compatibility at $t = 2$. As it is standard, incentive compatibility requires that, for any $\theta_1 \in \Theta_1$ and any $\xi \in \Xi$,

$$U_2(\theta_1, \xi) = U_2(\theta_1, \underline{\xi}) + \int_{\underline{\xi}}^{\xi} \frac{\partial v(\theta_1, \tilde{\xi})}{\partial \tilde{\xi}} q_2(\theta_1, \tilde{\xi}) d\tilde{\xi}, \tag{20}$$

with $q_2(\theta_1, \cdot)$ nondecreasing in ξ. Using the envelope theorem, we then have that incentive compatibility at $t = 1$ requires that, for any $\theta_1 \in \Theta_1$,

$$U(\theta_1) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta_1} \left[q_1(s) + \delta \int_{\underline{\xi}}^{\bar{\xi}} \frac{\partial v(s, \xi)}{\partial s} q_2(s, \xi) dG(\xi) \right] ds \tag{21}$$

Abstracting from the monotonicity conditions, we then have that the seller's *relaxed problem* can be stated as follows.

$$\mathcal{P}^r : \begin{cases} \max_{q_1(\cdot),q_2(\cdot),U(\cdot),U_2(\cdot)} \int_{\underline{\theta}}^{\bar{\theta}} \{\theta_1 q_1(\theta_1) - C(q_1(\theta_1)) \\ +\delta \int_{\underline{\xi}}^{\bar{\xi}} [v(\theta_1,\xi)q_2(\theta_1,\xi) - C(q_2(\theta_1,\xi))]dG(\xi) - U(\theta_1)\}dF(\theta) \\ \text{s.t. (20), (21), } U(\theta) \geq 0 \text{ and } U_2(\theta_1,\xi) \geq 0 \ \forall(\theta_1,\xi) \end{cases}$$

Integrating (21) by parts, we then have that the objective function in $\mathcal{P}^r$ can be rewritten as

$$\int_{\underline{\theta}}^{\bar{\theta}} \{ \left(\theta_1 - \tfrac{1-F(\theta_1)}{f(\theta_1)}\right) q_1(\theta_1) - C(q_1(\theta_1)) + \\ +\delta \int_{\underline{\xi}}^{\bar{\xi}} \left[\left(v(\theta_1,\xi) - \frac{\partial v(\theta_1,\xi)}{\partial \theta_1}\tfrac{1-F(\theta_1)}{f(\theta_1)}\right) q_2(\theta_1,\xi) - C(q_2(\theta_1,\xi))\right] dG(\xi) - U(\underline{\theta})\}dF(\theta_1).$$

It is then immediate that the solution to the program $\mathcal{P}^r$ is independent of the participation and the incentive-compatibility constraints for $t = 2$ and is such that $U(\underline{\theta}) = 0$ and

$$q_1(\theta_1) = \max\left\{q^{FB}(\theta_1) - \tfrac{1-F(\theta_1)}{f(\theta_1)}; 0\right\}$$
$$q_2(\theta_1,\xi) = \max\left\{q^{FB}(\theta_1,\xi) - \tfrac{\partial v(\theta_1,\xi)}{\partial\theta_1}\tfrac{1-F(\theta_1)}{f(\theta_1)}; 0\right\}$$

Turning to the general case in which $T \geq 2$ we can then redefine $\mathcal{P}^r$ to denote the relaxed program in which the seller maximizes her period-1 expected payoff subject to the sole constraints that $U(\underline{\theta}) = 0$ and

$$\begin{aligned} U(\theta_1) &= U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta_1} q_1(s)ds \\ &+ \int_{\underline{\theta}}^{\theta_1} \left\{ \int_{\underline{\xi}_2}^{\bar{\xi}_2} \cdots \int_{\underline{\xi}_T}^{\bar{\xi}_T} \left[\delta \frac{\partial v_2(s,\xi_2)}{\partial s} q_2(s,\xi_2) + \cdots + \delta^T \frac{\partial v_T(s,\xi^T)}{\partial s} q_T(s,\xi^T)\right] dW_2 \right\} ds \end{aligned}$$

for any $\theta_1 \in \Theta_1$, where $dW_2 \equiv dG_2(\xi_2) \times \cdots \times dG_T(\xi_T)$. We then have the following result.

Proposition 3 *The solution to the relaxed program is given by the schedules*

$$q_1(\theta_1) = \max\left\{q^{FB}(\theta_1) - \tfrac{1-F(\theta_1)}{f(\theta_1)}; 0\right\} \\ q_t(\theta_1,\xi^t) = \max\left\{q^{FB}(\theta_1,\xi^t) - \tfrac{\partial v_t(\theta_1,\xi^t)}{\partial\theta_1}\tfrac{1-F(\theta_1)}{f(\theta_1)}; 0\right\}, \ t \geq 2. \tag{22}$$

I will come back in a moment to the conditions that guarantee that the solution to the relaxed program coincides with the optimal mechanism. For the moment, suppose this is the case. The optimal mechanism then has the following properties.

Corollary 1 *Assume the optimal mechanism coincides with the solution to the relaxed program. Then,*

(i) *the quantity schedules (as well as the players' payoffs) coincide with the ones the seller would offer if the shocks were jointly observed instead of being the buyer's private information;*

(ii) *distortions may either increase or decrease over time, depending on the dynamics of the sensitivity of the buyer's future valuations to his initial type;*

(iii) *distortions are independent of the shocks' distributions;*

(iv) *the seller may find it optimal to exclude a buyer for a few periods and then serve him again once his valuation has sufficiently improved;*

(v) *starting from $t = 2$, a high-valuation buyer may receive a smaller quantity than a low-valuation buyer;*

(vi) *unless $v_s(\theta_1, \xi^s)$ is independent of θ_1 for all $s \geq t$, or $\theta_1 = \bar{\theta}$, it is never optimal for the seller to transfer the ownership of the production technology to the buyer.*

Provided the optimal mechanism coincides with the solution to the relaxed program, then whether the agent possesses private information about the shocks or not is irrelevant for the dynamics of the optimal quantities as well as for expected payoffs. The only possible effect of the buyer's private information about the shocks is on the dynamics of prices. In fact, as standard with quasi-linear payoffs, while the expected transfers are uniquely determined, their dynamics are not. In particular, while with public shocks the seller can always ask the buyer to pay everything up-front in period one, when the shocks are the buyer's private information, it is key for truthful information revelation to have the buyer pay also in subsequent periods.[5]

Next, consider parts (ii)-(iii). To better appreciate these parts, consider the case in which θ_t follows an AR(1) process, as in (19). In this case,

$$q_t(\theta_1, \xi^t) = \max\left\{ q^{FB}(\theta_1, \xi^t) - \prod_{s=2}^{t} a_s \tfrac{1-F(\theta_1)}{f(\theta_1)}; 0 \right\}. \tag{23}$$

Whether distortions decrease or increase over time then depends on how the buyer's initial type θ_1 impacts his future valuations. If $a_s < 1$ for all s, then distortions decrease over time. If however, $a_s = 1$ for all s, as in the random walk case, or when θ_t follows a Brownian motion (possibly with drift), then distortions remain constant over time. In this case the optimal mechanism is very

[5] One may think that the result in part (i) is a direct implication of the fact that the shocks are independent of the buyer's initial type. As shown below, this is not correct. Also note that the result relies on the fact that the supports of the buyer's valuations over any two adjacent periods exhibit enough overlapping (in a sense that will be made clear below). This is always the case with the continuum, but not necessarily with finitely-many types.

simple. It's a menu of long-term contracts according to which in each period the seller provides the buyer a quantity equal to the *first-best*, net of a distortion that depends only on the buyer's first-period type. More generally, distortions may also be affected by the shocks that the buyer experiences over time, but are independent of the shocks' distributions.

Finally, consider parts (iv)-(vi). That the seller may find it optimal to exclude the buyer for certain periods is an immediate implications of the fact that the buyer's virtual valuation

$$v_t(\theta_1, \xi^t) - \frac{\partial v_t(\theta_1,\xi^t)}{\partial \theta_1} \frac{1-F(\theta_1)}{f(\theta_1)}$$

can become negative for a few periods and then turn positive again after a sequence of favorable shocks. Similarly, that a high-valuation buyer may receive a lower quantity than a low-valuation buyer is a direct consequence of the fact that the optimal mechanism exhibits *memory*: the quantity that the buyer receives at any point in time is a function not only of the buyer's current valuation but also of how his valuation evolved over time.[6] Finally, that it is essentially never optimal to transfer the ownership of the production technology to the buyer is a consequence of the fact that distortions do not disappear over time. Unless the buyer's type in period one is the highest, or unless after a certain point in time valuations become independent of the buyer's initial type, then it is optimal for the seller to maintain distortions throughout the entire relationship. This immediately precludes the possibility of transferring ownership of the production technology to the buyer, for in that case the buyer implements the first best thereafter.

I now turn to the conditions that guarantee that the optimal schedules coincide with those that solve the relaxed program. First consider the following enlargement of the message space. For any $t \geq 2$, let $\underline{\underline{\xi}}_t$ and $\bar{\bar{\xi}}_t$ be recursively defined by

$$v_t(\bar{\theta}_1, \bar{\bar{\xi}}_2, ..., \bar{\bar{\xi}}_{t-1}, \underline{\underline{\xi}}_t) = \min \Theta_t \text{ and } v_t(\underline{\theta}_1, \underline{\underline{\xi}}_1, ..., \underline{\underline{\xi}}_{t-1}, \bar{\bar{\xi}}_t) = \max \Theta_t.$$

Then let

$$\tilde{\Xi}_t = [\underline{\underline{\xi}}_t, \bar{\bar{\xi}}_t]$$

This enlargement of the message space is redundant when $Supp[G_t]$ is unbounded; when instead it is bounded, it is a simple trick that permits us to describe in a convenient way the buyer's behavior off equilibrium. In fact, note that $\tilde{\Xi}_t$ is constructed so that the buyer can reveal his true valuation in period t, for any possible history of past reports (i.e. even after misreporting the shocks experienced in the past). Clearly in equilibrium, the only messages that the buyer will send are

[6] That the optimal contract exhibits memory is a property that holds true also in models with a binary type space. However,in those models high-valuation buyers always receive at least as much quantity as low valuation ones.

those that belong to the intervals $\Xi_t \equiv Supp[G_t]$.

Now suppose the following condition holds (all subsequent conditions are meant to hold over the extended state space $\Theta \times \tilde{\Xi}_1 \times \cdots \times \tilde{\Xi}_T$.)

Condition 1 (Markov property) *For any t, any $s > t$, and any $(\xi_{t+1}, ..., \xi_s)$,*

$$v_s(\theta_1, \xi^t, \xi_{t+1}, ..., \xi_s) = v_s(\tilde{\theta}_1, \tilde{\xi}^t, \xi_{t+1}, ..., \xi_s)$$

for any pair of histories (θ_1, ξ^t) and $(\tilde{\theta}_1, \tilde{\xi}^t)$ such that $v_t(\theta_1, \xi^t) = v_t(\tilde{\theta}_1, \tilde{\xi}^t)$.

Together with the assumption of independent shocks, this condition implies that the buyer's valuation follows a Markov process: for any t and any history of valuations $h_t \equiv (\theta_1, ..., \theta_t)$, the distribution of $(\theta_{t+1}, \theta_{t+2}, ..., \theta_T)$ given h_t depends only on θ_t.

Next, consider the following property.

Condition 2 (Strong monotonicity) *The quantity schedules $q_t(\cdot)$ are nondecreasing in each argument and satisfy the property that, for any $t \geq 2$, any $s \geq t+1$, any $(\xi_{t+2}, ..., \xi_s)$, and any pair $(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1})$ and $(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \tilde{\xi}_{t+1})$ such that $v_{t+1}(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1}) = v_{t+1}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \tilde{\xi}_{t+1})$,*[7]

$$q_s(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1}, \xi_{t+2}, ..., \xi_s) \leq (\textit{resp.} \geq)\ q_s(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \tilde{\xi}_{t+1}, \xi_{t+2}, ..., \xi_s)$$

if and only if $\xi_t \leq \tilde{\xi}_t$ (resp. $\xi_t \geq \tilde{\xi}_t$); similarly, for any (θ_1, ξ_2) and $(\tilde{\theta}_1, \tilde{\xi}_2)$ such that $v_2(\theta_1, \xi_2) = v_2(\tilde{\theta}_1, \tilde{\xi}_2)$ and for any $(\xi_3, ..., \xi_s)$, $q_s(\theta_1, \xi_2, ..., \xi_s) \leq$ (resp. $\geq$) $q_s(\tilde{\theta}_1, \tilde{\xi}_2, \xi_3, ..., \xi_s)$ if and only if $\theta_1 \leq \tilde{\theta}_1$ (resp. $\theta_1 \geq \tilde{\theta}_1$).

Together with the Markov property, this condition guarantees that, holding constant the buyer's valuation from period one to period $t-1$ and from period $t+1$ to period s, the quantity the seller provides in period s is higher the higher the shock (and hence the buyer's valuation) in period t. In other words, an unfavorable shock in period t followed by a favorable one in period $t+1$ leads to smaller future quantities than a favorable shock in period t followed by an unfavorable one in period $t+1$ that result in the same period $t+1$ valuation.

Note that the schedules that solve the relaxed program always satisfy the monotonicity properties of **Condition (2)** when the agent's valuation follows an AR(1) process—as defined in (19). By implication, these conditions are also satisfied when the agent's valuation evolves in continuous time following a Brownian motion. Furthermore, these conditions are also satisfied for example in

[7] If $t = 2$, drop ξ^{t-1} from all expressions. Similarly, if $s = t+1$, then drop $(\xi_{t+2}, ..., \xi_s)$.

the case of *multiplicative shocks*,[8] i.e. when

$$v_t(\theta_1, \xi^t) = \theta_1 \times \xi_2 \times \cdots \times \xi_t \tag{24}$$

with $Supp[G_s] \subseteq \mathbb{R}_+$, for any s. In this case, the schedules that solve the relaxed program are

$$q_t(\theta_1, \xi^t) = \max \left\{ (\xi_2 \times \cdots \times \xi_t) \left[\theta_1 - \tfrac{1-F(\theta_1)}{f(\theta_1)} \right] ; 0 \right\}.$$

We then have the following result.

Proposition 4 *Suppose the buyer's valuation follows a Markov process and that the schedules of Proposition 3 satisfy the strong monotonicity properties of* ***Condition (2)****. Then in any optimal mechanism the quantity schedules coincide with those in Proposition 3.*

The proof of Proposition 4 in the Appendix is in two steps. First, I show how starting from $t = T$ and proceeding backward one can construct recursively a profile of transfers that have the following property: if the buyer reported truthfully up to period $t-1$ he prefers reporting truthfully from period t onward rather than lying in period t and then reporting truthfully thereafter. In each period, these transfers are uniquely determined by the quantity schedules up to a scalar that can be set optimally to guarantee participation. The second step then shows how incentive compatibility can be established recursively, again starting from $t = T$. Using the trick of the enlarged message space, the best the buyer can do after lying in period t by reporting a shock $\hat{\xi}_t \neq \xi_t$ is to lie again in period $t+1$ by reporting a shock $\hat{\xi}^*_{t+1} = \hat{\xi}^*_{t+1}(\theta_1, \xi^{t-1}, \xi_t, \hat{\xi}_t, \xi_{t+1})$ implicitly defined by

$$v_{t+1}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \xi^*_{t+1}) = v_{t+1}(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1})$$

and then reporting truthfully from period $t+2$ onward. This property is an immediate consequence of the fact that the buyer's valuation follows a Markov process and of the fact that (by recursive construction) the mechanism is incentive compatible (on the equilibrium path) from period $t+1$ onward. Because the buyer's (equilibrium) continuation payoffs are increasing in the present and future level of trade, we then have that, when the quantity schedules satisfy the monotonicity properties of **Condition (2)**, the surplus the seller leaves to the buyer when the latter can lie only once is large enough to discourage him from lying also when he can misrepresent his private information multiple times.

The next proposition (which is technical and with little economics) identifies properties of the

[8] The case of multiplicative shocks seems particularly appropriate in the contest of borrower-lender relationships, in which case the shocks represent variations in the return to invested capital.

stochastic process that guarantee that indeed the schedules of Proposition 3 satisfy the monotonicity properties of **Condition (2)**.

Proposition 5 *Suppose that, in addition to the Markov property, the buyer's valuations satisfy the following conditions.*

(i) $v_t(\theta_1, \xi^t)$ *are concave in* θ_1 *and* $\frac{\partial^2 v_t(\theta_1,\xi^t)}{\partial\theta_1 \partial\xi_s} \leq 0$ *for any* (θ_1, ξ^t), $s \leq t$.

(ii) *For any* $t \geq 3$, *any* $s \geq t$ *and any* $(\theta_1, \xi^t, \xi_{t+1}, ..., \xi_s)$

$$\frac{\partial^2 v_s(\theta_1, \xi^t, \xi_{t+1}, ..., \xi_s)}{\partial\theta_1 \partial\xi_{t-1}} \frac{\partial v_t(\theta_1, \xi^t)}{\partial\xi_t} \leq \frac{\partial^2 v_s(\theta_1, \xi^t, \xi_{t+1}, ..., \xi_s)}{\partial\theta_1 \partial\xi_t} \frac{\partial v_t(\theta_1, \xi^t)}{\partial\xi_{t-1}};$$

Similarly, for $t = 2$ *and any* $(\theta_1, \xi_2, ..., \xi_s)$, $s \geq 2$,

$$\frac{\partial^2 v_s(\theta_1, \xi_2, ..., \xi_s)}{\partial\theta_1^2} \frac{\partial v_2(\theta_1, \xi_2)}{\partial\xi_2} \leq \frac{\partial^2 v_s(\theta_1, \xi_2, ..., \xi_s)}{\partial\theta_1 \partial\xi_2} \frac{\partial v_2(\theta_1, \xi^t)}{\partial\theta_1}.$$

Then the quantity schedules of Proposition 3 satisfy the monotonicity properties of ***Condition (2)****.*

Transitory shocks. I now turn to the case in which the shocks ξ_t have no persistent effect on the buyer's valuation. In this case, **Condition (1)** is violated—these processes are clearly not Markov. Nevertheless, essentially the same conditions that guarantee that the optimal mechanism coincides with the solution to the relaxed program in the case of a Markov process guarantee that the same holds true in the case of a stochastic process with transitory shocks.

Proposition 6 *Suppose there exist functions* $z_t : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *such that* $v_t(\theta_1, \xi^t) = z_t(\theta_1, \xi_t)$, *for any* (θ_1, ξ^t) *and any* t.

(i) *Assume the schedules of Proposition 3 are nondecreasing and, for any pair* (θ_1, ξ^t) *and* $(\tilde{\theta}_1, \tilde{\xi}^t)$ *such that* $v_t(\theta_1, \xi^t) = v_t(\tilde{\theta}_1, \tilde{\xi}^t)$,

$$q_t(\theta_1, \xi^t) \leq (\textit{resp.} \geq)\ q_t(\tilde{\theta}_1, \tilde{\xi}^t) \textit{ if and only if } \theta_1 \leq \tilde{\theta}_1\ (\textit{resp. } \theta_1 \geq \tilde{\theta}_1).$$

Then in any optimal mechanism the quantity schedules coincide with those in Proposition 3.

(ii) *Assume Condition (i) in Proposition 5 holds and, for any* $t \geq 2$,

$$\frac{\partial^2 v_t(\theta_1, \xi^t)}{\partial\theta_1^2} \frac{\partial v_t(\theta_1, \xi^t)}{\partial\xi_t} \leq \frac{\partial^2 v_t(\theta_1, \xi^t)}{\partial\theta_1 \partial\xi_t} \frac{\partial v_t(\theta_1, \xi^t)}{\partial\theta_1}.$$

Then the quantity schedules of Proposition 3 satisfy the monotonicity properties of part (i) of this proposition.

An example of a stochastic process with transitory shocks that satisfies the conditions in the preceding proposition is the following:

$$v_t\left(\theta_1, \xi^t\right) = \alpha\theta_1 + (1-\alpha)\xi_t$$

with $\alpha \in (0,1)$ and $Supp[G_t] = Supp[F]$ for all t. This is a stochastic process in which the set of possible types is stationary over time, i.e. $\Theta_t = \Theta_1$ for all t but which violates the Markov property. The dynamics of distortions is then extremely simple and given by

$$d_1(\theta_1) \equiv q_1^{FB}(\theta_1) - q_1(\theta_1) = \tfrac{1-F(\theta_1)}{f(\theta_1)}$$

while for any $t \geq 2$

$$d_t(\theta_1, \theta_2, ..., \theta_t) \equiv q_t^{FB}(\theta_1, \theta_2, ..., \theta_t) - q_t(\theta_1, \theta_2, ..., \theta_t) = \alpha\tfrac{1-F(\theta_1)}{f(\theta_1)}$$

Holding constant the buyer's valuation, distortions thus exhibit a downward discontinuity from period one to period two and are constant thereafter.

This example together with the random walk version of (19) thus suggests a possible explanation of what drives the results in Battaglini's two-type model. The property that distortions in the optimal contract vanish over time appears to be a consequence of the combination of two assumptions: that the buyer's valuation follows a Markov process and that the set of possible valuations is bounded and stationary over time. If the buyer's valuations follow a Markov process but the set of possible valuations is either unbounded or it changes over time, as in the random walk case, then distortions need not vanish. Likewise, if the set of possible valuations is bounded and stationary over time, but valuations do not follow a Markov process, as in the last example, then again there is no reason to expect distortions to disappear in the long-run.

The following observation is then an immediate implication of the preceding results.

Corollary 2 *Consider the following two processes:*

$$\textit{(a)}\ v_t(\theta_1, \xi^t) = a_t\theta_{t-1} + \xi_t = \prod_{j=2}^{t} a_j\theta_1 + \prod_{j=3}^{t} a_j\xi_2 + \cdots + a_t\xi_{t-1} + \xi_t$$

$$\textit{(b)}\ v_t(\theta_1, \xi^t) = \prod_{j=2}^{t} a_j\theta_1 + \xi_t$$

for some $(a_j)_{j=1}^T \in \mathbb{R}_+^T$. *The dynamics of distortions in any optimal mechanism are the same for (a) and (b).*

What differentiates the two processes in Corollary 2 is *only* the persistence of the shocks. Such a distinction, however, need not be relevant for the dynamics of distortions. The result holds more generally for any pair of processes for which the optimal mechanism coincides with the solution to the relaxed program and for which the dynamics of the sensitivity $\partial v_t(\theta_1, \xi^t)/\partial\theta_1$ of future valuations to the buyer's initial type is the same.

4.2 Finitely-many types

Consider now the same model with independent shocks examined above, but assume that in each period there are finitely many types. Continue to denote by $F(\theta_1)$ the c.d.f. of the buyer's first period type but now let $f(\theta_1)$ denote the corresponding probability distribution function instead of the density. Without loss, then order the buyer's first-period types so that $\theta_1^1 < \theta_1^2 < \cdots < \theta_1^N$ and, for any n, let $\Delta\theta^n \equiv \theta_1^{n+1} - \theta_1^n$ with $\Delta\theta^N \equiv 0$. Then for any $t \geq 2$ let

$$\Delta_{\theta^n} v_t(\theta_1^n, \xi^t) \equiv v_t(\theta_1^{n+1}, \xi^t) - v_t(\theta_1^n, \xi^t)$$

with again $\Delta_{\theta^N} v_t(\theta_1^n, \xi^t) \equiv 0$.

The support of the shocks' distributions G_t is now a finite set Ξ_t. As in the previous section, let $\tilde{\Xi}_t$ denote the extended message space defined recursively by

$$\tilde{\Xi}_t = \{\tilde{\xi}_t \in \mathbb{R} \text{ s.t. } v_t(\theta_1, \tilde{\xi}^{t-1}, \tilde{\xi}_t) = \theta_t \text{ for some } \theta_t \in \Theta_t \text{ and } (\theta_1, \tilde{\xi}^{t-1}) \in \Theta_t \times \tilde{\Xi}_t^{t-1}\}.$$

As in the previous section, these extended message spaces simply permit the buyer to reveal his type in period t after having lied in the past.

Now suppose the seller can observe the shocks (but not the buyer's initial type θ_1). Then consider the *relaxed program* in which the seller maximizes her expected payoff subject to the sole constraints that each type θ_1 must find it optimal to participate in the seller's mechanism in period one (but not necessarily thereafter) and truthfully reveal his type. Call this relaxed program $\mathcal{P}_r$. We then have the following result.

Lemma 3 *Suppose the schedules*

$$q_1(\theta_1^n) = \max\left\{q^{FB}(\theta_1^n) - \Delta\theta^n \left(\frac{1-F(\theta_1^n)}{f(\theta_1^n)}\right); 0\right\}$$
$$q_t(\theta_1^n, \xi^t) = \max\left\{q^{FB}(\theta_1^n, \xi^t) - \Delta_{\theta^n} v_t(\theta_1^n, \xi^t) \left(\frac{1-F(\theta_1^n)}{f(\theta_1^n)}\right); 0\right\}, \ t \geq 2$$

are increasing in each argument. Then they solve the relaxed program $\mathcal{P}_r$.

Now suppose the supports of the buyer's valuations overlap enough in the sense defined by the following condition.

Condition 3 (Overlapping supports) *Let* ξ_t *and* ξ'_t *be any two adjacent shocks for period* $t \geq 2$, *with* $\xi_t < \xi'_t$. *Then for any pair of adjacent shocks* ξ_{t+1} *and* ξ'_{t+1} *with* $\xi_{t+1} > \xi'_{t+1}$*and any* $s \geq t+1$

$$v_s(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1}, \xi_{t+2}, ..., \xi_s) = v_s(\theta_1, \xi^{t-1}, \xi'_t, \xi'_{t+1}, \xi_{t+2}, ..., \xi_s), \ \forall \ (\theta_1, \xi^{t-1}, \xi_{t+2}, ..., \xi_s)$$

Similarly, for any pair of adjacent period-1 types θ_1 *and* θ'_1, *with* $\theta_1 < \theta'_1$, *and any pair of adjacent shocks* ξ_2 *and* ξ'_2 *with* $\xi_2 > \xi'_2$, $v_s(\theta_1, \xi_2, \xi_3..., \xi_s) = v_2(\theta'_1, \xi'_2, \xi_3, ..., \xi_s)$ *for any* $s \geq 2$ *and any* $(\xi_3, ..., \xi_s)$.

When the overlapping condition holds, the result in Proposition 4 that the optimal schedules coincide with those that the seller would offer when the shocks are jointly observed continues to hold.

Proposition 7 *Suppose the buyer's valuation follows a finite Markov process with independent shocks and that the overlapping support condition holds. Assume the schedules in Lemma 3, extended over the enlarged state space* $\Theta \times \tilde{\Xi}^T$, *satisfy the strong monotonicity properties of* ***Condition 2****. Then in any optimal mechanism the quantity schedules coincide with those in Lemma 3 and each player obtains the same expected payoff as when the shocks to the buyer's valuation are jointly observed.*

An example of a stochastic process that satisfies the conditions in Proposition 7 is the process considered in Section 3: $v_t(\theta_1, \xi^t) = \theta_1 + \sum_{s=2}^{t} \xi_s$, with $\Theta_1 = \{\underline{\theta}, \bar{\theta}\}$, $\Xi_t = \{\xi_h, \xi_m, \xi_l\}$, $\xi_h = +\Delta\theta$, $\xi_m = 0$, $\xi_l = -\Delta\theta$, $\Pr(\theta_1 = \bar{\theta}) = v$, $\Pr(\xi_t = \xi_h) = x$ and $\Pr(\xi_t = \xi_m) = \alpha$, for all t. In this case, the optimal schedules are

$$q_t(\theta_1, \xi^t) = \begin{cases} q^{FB}(\theta_1, \xi^t) \text{ if } \theta_1 = \bar{\theta} \\ \max\left\{q^{FB}(\theta_1, \xi^t) - \frac{v}{1-v}\Delta\theta; 0\right\} \text{ if } \theta_1 = \underline{\theta} \end{cases}$$

In this example distortions in the contract for the low type are constant over time and have the familiar structure $d_t(\theta, \xi^t) = v\Delta\theta/(1-v)$ for all t.

This example also illustrates the role of the overlapping condition. For simplicity, assume $T = 2$. To induce the buyer to reveal truthfully the shock ξ_2, the seller can make all downward adjacent incentive-compatibility constraints for period two binding. This accounts to choosing prices $p_2(\theta_1, \xi)$ such that

$$U_2(\theta_1, \xi_h) - U_2(\theta_1, \xi_m) = \Delta\theta q_2(\theta_1, \xi_m) \tag{25}$$

$$U_2(\theta_1, \xi_m) - U_2(\theta_1, \xi_l) = \Delta\theta q_2(\theta_1, \xi_l) \tag{26}$$

for both $\theta_1 = \underline{\theta}$ and $\theta_1 = \bar{\theta}$. Lemma 1, then implies that a high type who mimics the low type in period one and then experiences a shock ξ_h strictly prefers to report the shock truthfully in period two than lying, whereas he is indifferent between reporting the true shock or one just below it when the shock he experiences is either ξ_m or ξ_l. This in turn implies that, when the transfers for period two satisfy (25) and (26), the expected surplus the seller must leave to the high type in period one is

$$U(\bar{\theta}) = U(\underline{\theta}) + \Delta\theta q_1(\underline{\theta}) + \Delta\theta \mathbb{E}_\xi \left[q_2(\underline{\theta}, \xi)\right]$$

This is exactly the same surplus the seller would give to the high type when shocks are jointly observed.

Now suppose instead that the shocks ξ_h, ξ_m and ξ_l are such that $\Delta\bar{\xi} \equiv \xi_h - \xi_m < \Delta\theta$ and $\Delta\underline{\xi} \equiv \xi_m - \xi_l < \Delta\theta$ (with $\Delta\xi = \Delta\bar{\xi} + \Delta\underline{\xi}$). From Lemma 1, we have that, when the high type lies in period one, then in period two, irrespective of the true shock ξ, he strictly prefers to report ξ_h than either ξ_m or ξ_l. By implication, the expected surplus the seller must leave to the high type in period one must be at least

$$U(\underline{\theta}) + \Delta\theta q_1(\underline{\theta}) + \mathbb{E}_\xi \left[U_2\left(\underline{\theta}, \xi_h\right) - U_2\left(\underline{\theta}, \xi\right) + \left(\Delta\theta + \xi - \xi_h\right) q_2(\underline{\theta}, \xi_h)\right]$$

which is strictly higher than the surplus $U(\underline{\theta}) + \Delta\theta q_1(\underline{\theta}) + \Delta\theta \mathbb{E}_\xi \left[q_2(\underline{\theta}, \xi)\right]$ that the seller must leave to the high type when shocks are jointly observed. There is then no reason to expect the schedules of Lemma 3 to be optimal any more. In fact, one can show that the best the seller can do in this example is to offer the following quantities to the low type[9] (assuming $q_2(\underline{\theta}, \xi)$ are monotonic in ξ which is always the case for v low enough)

$$\begin{aligned} q_1(\underline{\theta}) &= \max\{q^{FB}(\underline{\theta}) - \tfrac{\nu}{1-\nu}\Delta\theta; 0\} \\ q_2(\underline{\theta}, \xi_h) &= \max\{q^{FB}(\underline{\theta}, \xi_h) - \left(\tfrac{\nu}{1-\nu}\right)\left(\tfrac{\Delta\theta - \alpha\Delta\bar{\xi} - (1-\alpha-x)\Delta\underline{\xi}}{x}\right); 0\} \\ q_2(\underline{\theta}, \xi_m) &= \max\{q^{FB}(\underline{\theta}, \xi_m) - \left(\tfrac{\nu}{1-\nu}\right)\left(\tfrac{1-x}{\alpha}\right)\Delta\bar{\xi}; 0\} \\ q_2(\underline{\theta}, \xi_l) &= \max\{q^{FB}(\underline{\theta}, \xi_l) - \left(\tfrac{\nu}{1-\nu}\right)\Delta\underline{\xi}; 0\}. \end{aligned} \tag{27}$$

In contrast to the case of overlapping types, the distortions in the contract for the low-type now depend on the details of the shock distribution (i.e. on the probabilities x and α). However, as in the case of large shocks, it remains true that distortions need not vanish over time and need not be monotonic in the size of the shock to the buyer's valuation.

[9] The proof is available upon request.

5 Foundations for independent shocks

The results in the previous section have been established assuming the shocks to the buyer's valuation are independent. I now show that this assumption is actually less restrictive than it appears. The results in this section follow from arguments similar to those used to establish Lemmas 1 and 2 in Eso and Szentes (2007): the subsequent propositions extend their results to an arbitrary number of periods and an arbitrary number of decisions.

Proposition 8 *Suppose for any t and any $(\theta_1, ..., \theta_{t-1})$, the distribution of θ_t given $(\theta_1, ..., \theta_{t-1})$ is continuous and strictly increasing in θ_t. Then there exist a collection of real-valued functions $v_t : \mathbb{R}^t \longrightarrow \mathbb{R}$ and a collection of random variables ξ_t such that, for any $t \geq 2$, the process can be described by*

$$\theta_t = v_t(\theta_1, \xi^t)$$

with (θ_1, ξ^T) jointly independent.

Assuming independent shocks is thus truly without loss of generality in the case of a continuous process. The approach indicated in the previous section to represent the necessary conditions for incentive-compatibility (and the corresponding solution to the relaxed program) are thus quite general. On the other hand, the sufficient conditions of Proposition 4 require that, in the case of a Markov process, the functions v_t be increasing in each argument. The next proposition shows that this is equivalent to assuming *first-order-stochastic-dominance.*

Proposition 9 ***(i)*** *Assume the buyer's valuation follows a Markov process and that the conditional distribution function of θ_t given θ_{t-1} is continuous, strictly increasing in θ_t and strictly decreasing in θ_{t-1}. Then the corresponding functions v_t are strictly increasing in each argument.*

(ii) *Assume the process for θ_t satisfies the conditions in part (i) in this proposition. Then any pair of collections $(v_t, \xi_t)_{t=2}^T$ and $(\tilde{v}_t, \tilde{\xi}_t)_{t=2}^T$ that represent the same process is such that*

$$\frac{\partial v_t(\theta_1, \xi^t)}{\partial \theta_1} = \frac{\partial \tilde{v}_t(\theta_1, \tilde{\xi}^t)}{\partial \theta_1}$$

for any (θ_1, ξ^t) and $(\theta_1, \tilde{\xi}^t)$ for which $v_s(\theta_1, \xi^s) = \tilde{v}_s(\theta_1, \tilde{\xi}^s)$ for any $s \leq t$.

Part (i) is self-explanatory. Part (ii) establishes that, although the representation of a stochastic process by means of a collection of independent innovations is not unique, any representation of the same Markov process leads to the same solution to the relaxed program. Using the aforementioned results, the sufficient conditions of Proposition 5—which guarantee that the solution to the relaxed program coincides with the optimal mechanism—can then be translated in terms of conditional distributions.

Proposition 10 *Assume the buyer's valuation follows a Markov process and, for any $t \geq 2$, let $\mathcal{F}_t(\theta_t; \theta_{t-1})$ denote the conditional distribution of θ_t given θ_{t-1} with $f_t(\theta_t; \theta_{t-1})$ denoting the corresponding density. Suppose the functions $\mathcal{F}_t$ satisfy the properties of part (i) in Proposition 9 and that, in addition,*

$$[\partial \mathcal{F}_t(\theta_t; \theta_{t-1})/\partial \theta_{t-1}] \,/\, [f_t(\theta_t; \theta_{t-1})]$$

are well-defined, and increasing in both θ_t and θ_{t-1}. Then the corresponding functions v_t satisfy the conditions of Proposition 5.

Similarly, one can show that the results in Proposition 6 for the case of *transitory* shocks extend more generally to any stochastic process that satisfies the property that the conditional distribution of θ_t given $(\theta_1, ..., \theta_{t-1})$ is strictly increasing in θ_t, decreasing in θ_1, and is independent of $(\theta_2, ..., \theta_{t-1})$.

The results in the previous section for the case of independent shocks thus apply to a quite rich class of continuous stochastic processes.

6 Appendix

Proof of Lemma 1. Because $\phi_2(\theta_1, \xi'')$ and $\phi_2(\theta_1, \xi')$ are incentive-compatible, there exists a tariff $\mathcal{T}_2 : \mathbb{R}_+ \to \mathbb{R}$ such that $\mathcal{T}_2(q_2(\theta_1, \xi'')) = p_2(\theta_1, \xi'')$, $\mathcal{T}_2(q_2(\theta_1, \xi')) = p_2(\theta_1, \xi')$ and $\mathcal{T}_2(q_2) = +\infty$ for any $q_2 \neq q_2(\theta_1, \xi'), q_2(\theta_1, \xi'')$. Furthermore, necessarily $q_2(\theta_1, \xi') < q_2(\theta_1, \xi'')$. The result then follows directly from Topkis' Monotonicity Theorem using the fact that the function $g(\theta_2, q_2) \equiv \theta_2 q_2$ satisfies the increasing difference property. ■

Proof of Proposition 1. I prove the result by showing that, whenever the quantities $q_2(\underline{\theta}, \xi)$ characterized in Lemma 2 are nondecreasing in ξ, there exists a profile of transfers that along with the quantities of Lemma 2 satisfies all the constraints in $\mathcal{P}$.

Start with $t = 2$. Take any profile of transfers $p_2(\theta, \xi)$ such that $U_2(\theta, \xi_l) \geq 0$ for $\theta = \bar{\theta}, \underline{\theta}$, (16) are saturated and

$$\begin{aligned} U_2\left(\bar{\theta},\xi_h\right)-U_2\left(\bar{\theta},\xi_m\right) &= \Delta\bar{\xi}q_2\left(\theta,\xi_h\right) \\ U_2\left(\bar{\theta},\xi_m\right)-U_2\left(\bar{\theta},\xi_l\right) &= \Delta\underline{\xi}q_2\left(\theta,\xi_m\right) \end{aligned} \tag{28}$$

The constraints (IR-2) are then trivially satisfied. The monotonicity of $q_2(\theta, \xi)$ in ξ along with Lemma 1 then guarantee that the mechanism is incentive-compatible at $t = 2$ i.e. that $U_2(\theta, \xi) = \tilde{U}_2(\theta, \xi; \theta)$ for any (θ, ξ).

Next, consider the agent's incentives at $t = 1$. First consider a type $\bar{\theta}$ who reports $\underline{\theta}$ at $t = 1$. The fact that the constraints (16) bind together with Condition (17) and Lemma 1 implies that the best $\bar{\theta}$ can do at $t = 2$ after lying at $t = 1$ is to report the shock ξ truthfully. Constraint (15) then guarantees that $\bar{\theta}$ obtains his maximal intertemporal payoff by reporting truthfully in both periods, i.e. that (IC) is satisfied for $\bar{\theta}$.

Now consider the incentives of a low type. The fact that constraints (28) bind together with Condition (17) and Lemma 1 imply that the best $\underline{\theta}$ can do at $t = 2$ after lying at $t = 1$ is to report the shock ξ truthfully. The following is then a sufficient condition for (IC) to be satisfied for $\underline{\theta}$:

$$\begin{aligned} U(\underline{\theta}) \geq\ & U(\bar{\theta}) - \Delta\theta q_1(\bar{\theta}) - \delta\{\overline{x}U_2(\bar{\theta},\xi_h) + \overline{\alpha}U_2(\bar{\theta},\xi_m) + (1-\overline{\alpha}-\overline{x})U_2(\bar{\theta},\xi_l)\} \\ &+ \delta\{\underline{x}[U_2(\bar{\theta},\xi_h) - \Delta\theta q_2(\bar{\theta},\xi_h)] + \underline{\alpha}[U_2(\bar{\theta},\xi_m) - \Delta\theta q_2(\bar{\theta},\xi_m)] \\ &+ (1-\underline{\alpha}-\underline{x})[U_2(\bar{\theta},\xi_l) - \Delta\theta q_2(\bar{\theta},\xi_l)]\} \end{aligned} \tag{29}$$

The right hand side in (29) is the payoff that $\underline{\theta}$ obtains by reporting $\bar{\theta}$ at $t = 1$ and then truthfully announcing the shock ξ at $t = 2$. Condition (29) can be conveniently rewritten as

$$\begin{aligned} U(\bar{\theta}) \leq\ & U(\underline{\theta}) + \Delta\theta q_1(\bar{\theta}) + \delta\{(\overline{x}-\underline{x})[U_2(\bar{\theta},\xi_h) - U_2(\bar{\theta},\xi_m)] \\ &+ (\bar{x}+\bar{\alpha}-\underline{\alpha}-\underline{x})[U_2(\bar{\theta},\xi_m) - U_2(\bar{\theta},\xi_l)] + \underline{x}\Delta\theta q_2(\bar{\theta},\xi_h) \\ &+ \underline{\alpha}\Delta\theta q_2(\bar{\theta},\xi_m) + (1-\underline{\alpha}-\underline{x})\Delta\theta q_2(\bar{\theta},\xi_l)\} \end{aligned} \tag{30}$$

The fact that (15), (16) and (28) are saturated together with Condition (17) and the fact that $q_2(\underline{\theta}, \xi) \leq q_2(\bar{\theta}, \xi)$ for all ξ then implies that (30) is satisfied. I conclude that (IC) is also satisfied for $\underline{\theta}$. Finally, that (IR-1) is satisfied for $\bar{\theta}$ follows immediately from (15) and the fact that $U(\underline{\theta}) = 0$. ■

Proof of Proposition 4. I want to show that there exists a mechanism that implements the allocations of Proposition 3 and gives each player the same expected payoff as the solution to the relaxed program. I prove the result in two steps. Step 1 constructs the mechanism. Step 2 then proves that, when the quantity schedules satisfy the strong monotonicity properties of **Condition**

(2), a buyer who reported truthfully up to period $t-1$, (weakly) prefers to report truthfully from period t onward rather than lying in period t and then choosing *optimally what to report in any subsequent period.*

Step 1. Consider the following mechanism. For each t, the quantity schedules $q_t(\theta_1, \xi^t)$ are as in Proposition 3, defined over the enlarged state space $\Theta_1 \times \tilde{\Xi}^T$. The transfer schedules are constructed as follows. Start from $t = T$. Given the schedules $q_T(\cdot)$, take any profile of transfers $p_T(\cdot)$ such that, for any $(\theta_1, \xi^{T-1}, \xi_T)$,

$$p_T(\theta_1, \xi^{T-1}, \xi_T) = v_T(\theta_1, \xi^{T-1}, \xi_T) q_T(\theta_1, \xi^{T-1}, \xi_T) - U_T(\theta_1, \xi^{T-1}, \xi_T)$$

with

$$U_T(\theta_1, \xi^{T-1}, \xi_T) = U_T(\theta_1, \xi^{T-1}, \underline{\underline{\xi}}_T) + \int_{\underline{\underline{\xi}}_t}^{\xi_T} \frac{\partial v_T(\theta_1, \xi^{T-1}, \tilde{\xi}_T)}{\partial \tilde{\xi}_T} q_T(\theta_1, \xi^{T-1}, \tilde{\xi}_T) d\tilde{\xi}_T \tag{31}$$

for an arbitrary scalar $U_T(\theta_1, \xi^{T-1}, \underline{\underline{\xi}}_T) \geq 0$. Proceeding backward, starting from $T-1$, then consider an arbitrary $t < T$. Holding constant the transfers constructed for periods $t+1$ onward, take any profile of transfers $p_t(\cdot)$ such that

$$p_t(\theta_1, \xi^{t-1}, \xi_t) = v_t(\theta_1, \xi^{t-1}, \xi_t) q_t(\theta_1, \xi^{t-1}, \xi_t) - U_t(\theta_1, \xi^{t-1}, \xi_t)$$

where $U_t(\theta_1, \xi^{t-1}, \underline{\underline{\xi}}_t)$ is an arbitrary non-negative scalar whereas, for any $\xi_t > \underline{\underline{\xi}}_t$,

$$\begin{aligned} U_t(\theta_1, \xi^{t-1}, \xi_t) \quad = \quad & U_t(\theta_1, \xi^{t-1}, \underline{\underline{\xi}}_t) + \int_{\underline{\underline{\xi}}_t}^{\xi_t} \Bigg\{ \frac{\partial v_t(\theta_1, \xi^{t-1}, \tilde{\xi}_t)}{\partial \tilde{\xi}_t} q_t(\theta_1, \xi^{t-1}, \tilde{\xi}_t) \\ & + \int_{\underline{\xi}_{t+1}}^{\bar{\xi}_{t+1}} \cdots \int_{\underline{\xi}_T}^{\bar{\xi}_T} [\delta \frac{\partial v_{t+1}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1})}{\partial \tilde{\xi}_t} q_{t+1}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}) \\ & + \cdots \\ & + \delta^{T-t} \frac{\partial v_T(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}, ..., \xi_T)}{\partial \tilde{\xi}_t} q_T(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}, ..., \xi_T)] dW_t \Bigg\} d\tilde{\xi}_t \end{aligned} \tag{32}$$

where $dW_t \equiv dG_{t+1}(\xi_{+1}) \times \cdots \times dG_T(\xi_T)$.

Step 2. I now want to show that the mechanism constructed in Step 1 guarantees that, a buyer who reported truthfully up to period $t-1$, (weakly) prefers to report truthfully from period t onwards rather than lying in period t and then *choosing optimally what to report at any subsequent date.* I establish the result by backward induction.

Start from $t = T$. Because the schedules $q_T(\cdot)$ are nondecreasing in $\tilde{\xi}_T$, standard results in static mechanism design—e.g. Myerson (1981) and Fudenberg and Tirole (1991)—imply that, a buyer who reported (θ_1, ξ^{T-1}) in the past and whose period-T valuation is $v_T(\theta_1, \xi^{T-1}, \xi_T)$ prefers to report ξ_T than reporting any other ξ'_T, for any $\xi_T, \xi'_T \in \tilde{\Xi}_T$. The fact that the buyer's valuation follows a Markov process then also implies, a buyer who reported (θ_1, ξ^{T-1}) in the past and who experiences a shock ξ_T in period T finds it optimal to report the shock ξ_T truthfully when the true shocks he experienced in the past are $(\tilde{\theta}_1, \tilde{\xi}^{T-1})$, for any $(\tilde{\theta}_1, \tilde{\xi}^{T-1})$ such that $v_{T-1}(\tilde{\theta}_1, \tilde{\xi}^{T-1}) = v_{T-1}(\theta_1, \xi^{T-1})$. That a buyer who reported $\left(\theta_1, \xi^{T-1}\right)$ in the past and whose period-T valuation is $v_T(\theta_1, \xi^{T-1}, \xi_T)$ finds it optimal to participate in period T follows from (31) along with the fact that $U_T(\theta_1, \xi^{T-1}, \underline{\xi}_T) \geq 0$.

Now take any $t < T$ and suppose the result established for period T holds more generally for any $s > t$. If the buyer reported truthfully in the past and reports truthfully in period t, then he best he can do is to report truthfully also at any subsequent period, in which case his expected payoff is $U_t(\theta_1, \xi^{t-1}, \xi_t)$. If instead he reports $\hat{\xi}_t \neq \xi_t$ in period t, he obtains

$$v_t(\theta_1, \xi^{t-1}, \xi_t) q_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) - p_t(\theta_1, \xi^{t-1}, \hat{\xi}_t)$$

in period t and then his maximal continuation payoff starting from period $t+1$ is

$$\tilde{U}_{t+1}\left(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1}; \theta_1, \xi^{t-1}, \hat{\xi}_t\right).$$

Incentive compatibility then requires that, for any $\left(\theta_1, \xi^{t-1}, \xi_t\right)$ and any $\hat{\xi}_t \neq \xi_t$,

$$U_t(\theta_1, \xi^{t-1}, \xi_t) \geq \hat{U}_t(\theta_1, \xi^{t-1}, \xi_t; \theta_1, \xi^{t-1}, \hat{\xi}_t) \tag{33}$$

where

$$\begin{aligned} \hat{U}_t(\theta_1, \xi^{t-1}, \xi_t; \theta_1, \xi^{t-1}, \hat{\xi}_t) \quad \equiv \quad & v_t(\theta_1, \xi^{t-1}, \xi_t) q_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) - p_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) \\ & + \delta \int_{\underline{\xi}_{t+1}}^{\bar{\xi}_{t+1}} \tilde{U}_{t+1}\left(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1}; \theta_1, \xi^{t-1}, \hat{\xi}_t\right) dG_{t+1}(\xi_{t+1}) \end{aligned}$$

Now note that

$$\begin{aligned} & \hat{U}_t(\theta_1, \xi^{t-1}, \xi_t; \theta_1, \xi^{t-1}, \hat{\xi}_t) \\ = \quad & U_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) + [v_t(\theta_1, \xi^{t-1}, \xi_t) - v_t(\theta_1, \xi^{t-1}, \hat{\xi}_t)] q_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) \\ & + \delta \int_{\underline{\xi}_{t+1}}^{\bar{\xi}_{t+1}} \left\{ \tilde{U}_{t+1}\left(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1}; \theta_1, \xi^{t-1}, \hat{\xi}_t\right) - U_{t+1}\left(\theta_1, \xi^{t-1}, \hat{\xi}_t, \xi_{t+1}\right) \right\} dG_{t+1}(\xi_{t+1}). \end{aligned}$$

Because $U_{t+1}\left(\theta_1, \xi^{t-1}, \hat{\xi}_t, \xi_{t+1}\right) = \tilde{U}_{t+1}\left(\theta_1, \xi^{t-1}, \hat{\xi}_t, \xi_{t+1}; \theta_1, \xi^{t-1}, \hat{\xi}_t\right)$, the constraint (33) can be rewritten as

$$
\begin{aligned}
&U_t(\theta_1, \xi^{t-1}, \xi_t) - U_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) \geq \\
&\int_{\hat{\xi}_t}^{\xi_t} \left[\frac{\partial v_t(\theta_1, \xi^{t-1}, \tilde{\xi}_t)}{\partial \tilde{\xi}_t} q_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) \right] d\tilde{\xi}_t \\
&+\delta \int_{\hat{\xi}_t}^{\xi_t} \left\{ \int_{\underline{\xi}_{t+1}}^{\bar{\xi}_{t+1}} \left[\frac{\partial \tilde{U}_{t+1}\left(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}; \theta_1, \xi^{t-1}, \hat{\xi}_t\right)}{\partial \tilde{\xi}_t} \right] dG_{t+1}(\xi_{t+1}) \right\} d\tilde{\xi}_t
\end{aligned} \tag{34}
$$

where, using (32)

$$
\begin{aligned}
& U_t(\theta_1, \xi^{t-1}, \xi_t) - U_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) \\
= & \int_{\hat{\xi}_t}^{\xi_t} \left\{ \frac{\partial v_t(\theta_1, \xi^{t-1}, \tilde{\xi}_t)}{\partial \tilde{\xi}_t} q_t(\theta_1, \xi^{t-1}, \tilde{\xi}_t) \right. \\
& + \int_{\underline{\xi}_{t+1}}^{\bar{\xi}_{t+1}} \cdots \int_{\underline{\xi}_T}^{\bar{\xi}_T} [\delta \frac{\partial v_{t+1}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1})}{\partial \tilde{\xi}_t} q_{t+1}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}) \\
& + \cdots \\
& \left. +\delta^{T-t} \frac{\partial v_T(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}, ..., \xi_T)}{\partial \tilde{\xi}_t} q_T(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}, ..., \xi_T)] dW_t \right\} d\tilde{\xi}_t.
\end{aligned} \tag{35}
$$

Furthermore, using the envelope theorem,

$$
\begin{aligned}
& \frac{\partial \tilde{U}_{t+1}\left(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}; \theta_1, \xi^{t-1}, \hat{\xi}_t\right)}{\partial \tilde{\xi}_t} \\
= & \frac{\partial v_{t+1}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1})}{\partial \tilde{\xi}_t} q_{t+1}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \hat{\xi}^*_{t+1}) \\
& +\delta \int_{\underline{\xi}_{t+2}}^{\bar{\xi}_{t+2}} \frac{\partial \tilde{U}_{t+2}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}, \xi_{t+2}; \theta_1, \xi^{t-1}, \hat{\xi}_t, \hat{\xi}^*_{t+1})}{\partial \tilde{\xi}_t} dG_{t+2}(\xi_{t+2})
\end{aligned}
$$

where $\hat{\xi}^*_{t+1} = \hat{\xi}^*_{t+1}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \hat{\xi}_t, \xi_{t+1})$ is the buyer's optimal report for period $t+1$. The fact that the buyer's valuation follows a Markov process along with the fact that the mechanism is incentive-compatible (on the equilibrium path) from period $t+1$ onward, implies that the best the buyer can do after lying in period t is to lie again in period $t+1$ by reporting a shock ξ^*_{t+1} implicitly defined by

$$
v_{t+1}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \xi^*_{t+1}) = v_{t+1}(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1})
$$

and then reporting truthfully thereafter. This means that the right hand side in (34) is given by

$$\int_{\hat{\xi}_t}^{\xi_t} \Bigg\{ \frac{\partial v_t(\theta_1, \xi^{t-1}, \tilde{\xi}_t)}{\partial \tilde{\xi}_t} q_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) \tag{36}$$

$$+ \int_{\underline{\xi}_{t+1}}^{\bar{\xi}_{t+1}} \cdots \int_{\underline{\xi}_T}^{\bar{\xi}_T} [\delta \frac{\partial v_{t+1}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1})}{\partial \tilde{\xi}_t} q_{t+1}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \hat{\xi}^*_{t+1})$$

$$+\delta^2 \frac{\partial v_{t+2}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}, \xi_{t+2})}{\partial \tilde{\xi}_t} q_{t+2}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \hat{\xi}^*_{t+1}, \xi_{t+2})$$

$$+ \cdots +$$

$$+\delta^{T-t} \frac{\partial v_T(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}, ..., \xi_T)}{\partial \tilde{\xi}_t} q_T(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \hat{\xi}^*_{t+1}, \xi_{t+2} ..., \xi_T)] dW_t \Bigg\} d\tilde{\xi}_t$$

Combining (35) with (36), it is then immediate to see that the inequality in (34) holds whenever the quantity schedules satisfy the strong monotonicity properties of **Condition (2)**. Finally, that a buyer who reported (θ_1, ξ^{T-1}) in the past and whose period-t valuation is $v_t(\theta_1, \xi^{T-1}, \xi_t)$ finds it optimal to participate in period t follows from (32) together with the fact that $U_t(\theta_1, \xi^{t-1}, \underline{\xi}_t) \geq 0$.

Iterating up to $t = 1$, then proves that the mechanism constructed in step 1 is indeed incentive compatible and it induces the buyer to participate in each period. ■

Proof of Proposition 5. Condition (i) in the proposition, together with the monotone hazard rate assumption, guarantees that the quantity schedules in (22) are non-decreasing. Now take any $t \geq 3$ and any $(\theta_1, \xi^{t-2}, \xi_{t-1}, \xi_t)$ and $(\theta_1, \xi^{t-2}, \tilde{\xi}_{t-1}, \tilde{\xi}_t)$ such that[10]

$$v_t\left(\theta_1, \xi^{t-2}, \xi_{t-1}, \xi_t\right) = v_t\left(\theta_1, \xi^{t-2}, \tilde{\xi}_{t-1}, \tilde{\xi}_t\right) \tag{37}$$

Without loss, assume $\xi_{t-1} \leq \tilde{\xi}_{t-1}$. The result for $\xi_{t-1} > \tilde{\xi}_{t-1}$ follows from similar arguments. Clearly, if

$$\max \left\{ v_s(\theta_1, \xi^{t-2}, \xi_{t-1}, \xi_t, \xi_{t+1}, ..., \xi_s) - \frac{\partial v_s(\theta_1, \xi^{t-2}, \xi_{t-1}, \xi_t, \xi_{t+1}, ..., \xi_s)}{\partial \theta_1} \frac{1-F(\theta_1)}{f(\theta_1)}; 0 \right\} = 0$$

then necessarily

$$q_s(\theta_1, \xi^{t-2}, \tilde{\xi}_{t-1}, \tilde{\xi}_t, \xi_{t+1}, ..., \xi_s) \geq q_s(\theta_1, \xi^{t-2}, \xi_{t-1}, \xi_t, \xi_{t+1}, ..., \xi_s) \tag{38}$$

[10] If $t = 3$, then drop ξ^{t-2} from all expressions.

in which case the strong monotonicity property of **Condition (2)** is satisfied. Thus assume

$$v_s(\theta_1, \xi^{t-2}, \xi_{t-1}, \xi_t, \xi_{t+1}, ..., \xi_s) - \frac{\partial v_s(\theta_1, \xi^{t-2}, \xi_{t-1}, \xi_t, \xi_{t+1}, ..., \xi_s)}{\partial \theta_1} \frac{1-F(\theta_1)}{f(\theta_1)} > 0$$

Condition (1) then implies that (38) is satisfied if and only if

$$\frac{\partial v_s(\theta_1, \xi^{t-2}, \tilde{\xi}_{t-1}, \tilde{\xi}_t, \xi_{t+1}, ..., \xi_s)}{\partial \theta_1} \leq \frac{\partial v_s(\theta_1, \xi^{t-2}, \xi_{t-1}, \xi_t, \xi_{t+1}, ..., \xi_s)}{\partial \theta_1} \tag{39}$$

Now, using the Implicit Function theorem applied to (37), there exists a function $\xi_t^* : \Xi_{t-1} \longrightarrow \tilde{\Xi}_t$ such that, for any $x \in \Xi_{t-1}$,

$$v_t\left(\theta_1, \xi^{t-2}, x, \xi_t^*(x)\right) = v_t\left(\theta_1, \xi^{t-2}, \xi_{t-1}, \xi_t\right)$$

with

$$\frac{d\xi_t^*(x)}{dx} = -\frac{\frac{\partial v_t\left(\theta_1, \xi^{t-2}, x, \xi_t^*(x)\right)}{\partial \xi_{t-1}}}{\frac{\partial v_t\left(\theta_1, \xi^{t-2}, x, \xi_t^*(x)\right)}{\partial \xi_t}} \tag{40}$$

It follows that

$$\begin{aligned} \frac{\partial v_s(\theta_1, \xi^{t-2}, \tilde{\xi}_{t-1}, \tilde{\xi}_t, \xi_{t+1}, ..., \xi_s)}{\partial \theta_1} \quad = \quad & \frac{\partial v_s(\theta_1, \xi^{t-2}, \xi_{t-1}, \xi_t, \xi_{t+1}, ..., \xi_s)}{\partial \theta_1} \\ & + \int_{\xi_{t-1}}^{\tilde{\xi}_{t-1}} \frac{\partial^2 v_s(\theta_1, \xi^{t-2}, x, \xi_t^*(x), \xi_{t+1}, ..., \xi_s)}{\partial \theta_1 \partial \xi_{t-1}} dx \\ & + \int_{\xi_{t-1}}^{\tilde{\xi}_{t-1}} \frac{\partial^2 v_s(\theta_1, \xi^{t-2}, x, \xi_t^*(x), \xi_{t+1}, ..., \xi_s)}{\partial \theta_1 \partial \xi_t} \frac{d\xi_t^*(x)}{dx} dx \end{aligned} \tag{41}$$

Substituting (40) into (41), we then have that (39) is satisfied if condition (ii) in the proposition holds. Similar arguments establish the result for $t = 2$. ■

Proof of Proposition 6. Consider the same mechanism constructed in Step 1 in Proposition 4. Although the Markov property is violated, the result follows from essentially the same steps as in the proof of Proposition 4 by noting that the buyer's value function (and hence his optimal behavior) in any period $t \geq 2$ depend only on the buyer's type in period one, on what he reported in period one, and on the shock experienced in period t. The mechanism constructed in Step 1 in Proposition 4 then guarantees that a buyer who lied in period one will lie again at any subsequent period by reporting a shock $\hat{\xi}_t = \hat{\xi}_t(\theta_1, \xi_2, ..., \xi_t; \hat{\theta}_1, \hat{\xi}_2, ..., \hat{\xi}_{t-1})$ defined by

$$z_t(\theta_1, \xi_t) = z_t(\hat{\theta}_1, \hat{\xi}_t)$$

On the contrary, a buyer who reported truthfully in period one and who lies in period t, will report truthfully from period $t+1$ thereafter. Following the same arguments as in the Proof of Proposition 4 it is then easy to see that when the schedules $q_t(\cdot)$ of Proposition 3 satisfy the conditions of Part (i) in the Proposition then the buyer finds it optimal to report truthfully at any point in time. Finally, following the same steps as in the Proof of Proposition 5 it is easy to see that when the conditions of Part (ii) hold, the schedules $q_t(\cdot)$ of Proposition 3 satisfy the conditions of Part (i). ■

Proof of Lemma 3. The result follows from two simple observations. First, to minimize the surplus the seller leaves to the buyer, it is optimal to make the participation constraint of the lowest type θ_1^1 and all downward adjacent incentive compatibility constraints binding. Maximizing the seller's expected revenue subject to these constraints gives the schedules in the lemma. Next, note that, because the schedules are monotonic in each argument, then the same monotone-comparative statics results used to establish Lemma 1 guarantee that all other incentive-compatibility constraints are satisfied. That all all types find it optimal to participate is immediate. ■

Proof of Proposition 7.

The proof is in two steps and parallels that of Proposition 4. Step 1 constructs a profile of transfers that, along with the schedules of Lemma 3—extended over the enlarged state space $\Theta \times \tilde{\Xi}^T$—gives the buyer the same expected payoff as the solution to the relaxed program (and, by implication, the seller the same expected profits). Step 2, shows that the mechanism constructed in Step 1 has the following properties: a buyer who reported truthfully up to period $t-1$ finds it optimal to participate in the seller's mechanism in period t and (weakly) prefers to report truthfully from period t onward rather than lying in period t and then choosing *optimally what to report in any subsequent period*. Together, these properties guarantee that the schedules that solve the relaxed program coincide with the optimal ones.

Step 1. Let $q_t(\theta_1, \xi^t)$ denote the quantity schedules of Lemma 3, extended over the enlarged state space $\Theta \times \tilde{\Xi}^T$. Next, consider the following transfers—again defined over the enlarged state space $\Theta \times \tilde{\Xi}^t$. Starting from period T and proceeding backward, the transfers $p_t(\theta_1, \xi^t)$ are constructed so that, for any $(\theta_1, \xi^{t-1}, \xi_t) \in \Theta \times \tilde{\Xi}^t$, a buyer who reported (θ_1, ξ^{t-1}) in the past and whose period-t valuation is $v_t(\theta_1, \xi^{t-1}, \xi_t)$ is *indifferent* between reporting truthfully from period t onward and lying in period t by reporting a shock immediately below the true shock ξ_t and then reporting truthfully thereafter. That is, the transfers $p_t(\theta_1, \xi^t)$ are constructed so that all downward adjacent local incentive-compatibility constraints bind. In addition, the transfers $p_t(\theta_1, \xi^t)$ satisfy the property that, after any $(\theta_1, \xi^{t-1}) \in \Theta \times \tilde{\Xi}^{t-1}$, type $v_t(\theta_1, \xi^{t-1}, \underline{\underline{\xi}}_t)$ obtains a zero expected continuation payoff.

Step 2. Now consider the buyer's incentives in period t after he has reported truthfully in all preceding periods. By construction, the transfers of Step 1 guarantee that the buyer is indifferent between reporting truthfully from period t onward and lying in period t by reporting a shock immediately below the true one and then reporting truthfully thereafter. I now want to show that, when the schedules of the relaxed program of Lemma 3 satisfy the monotonicity properties of **Condition (2)** and the overlapping condition holds, then the transfers of Step 1 also guarantee that the buyer prefers to report truthfully from period t onward rather than lying in period t and then choosing optimally what to report at any subsequent period.

First note that a buyer who reported truthfully up to period $T-1$ finds it optimal to report truthfully also in period T. This follows directly from Lemma 1 along with the fact that all downward adjacent incentive-compatibility constraints bind. The rest of the proof is establish by induction. Suppose the mechanism is incentive-compatible (on the equilibrium path) from period $t+1$ onward. As in the proof of Proposition 4, this implies that the best the buyer can do after reporting $\hat{\xi}_t \neq \xi_t$ in period t is to report a shock $\hat{\xi}^*_{t+1} = \hat{\xi}^*_{t+1}(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1}; \theta_1, \xi^{t-1}, \hat{\xi}_t)$ in period $t+1$ implicitly defined by

$$v_{t+1}(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1}) = v_{t+1}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \hat{\xi}^*_{t+1})$$

and then reporting truthfully in all subsequent periods. The continuation payoff the buyer obtains by reporting $\hat{\xi}_t$ in period t is thus equal to

$$\begin{aligned} & \hat{U}_t(\theta_1, \xi^{t-1}, \xi_t; \theta_1, \xi^{t-1}, \hat{\xi}_t) \qquad (42) \\ = \; & U_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) + \left[v_t(\theta_1, \xi^{t-1}, \xi_t) - v_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) \right] q_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) \\ & + \delta \mathbb{E}_{\xi_{t+1}} \left[U_{t+1}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \hat{\xi}^*_{t+1}(\theta_1, \xi^{t-1}, \xi_t, \xi_{t+1}; \theta_1, \xi^{t-1}, \hat{\xi}_t)) - U_{t+1}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \xi_{t+1}) \right] \end{aligned}$$

On the contrary, the continuation payoff the buyer obtains by reporting the shock ξ_t truthfully in period t is $U_t(\theta_1, \xi^{t-1}, \xi_t)$. Now, suppose $\hat{\xi}_t < \xi_t$ and, for any $\hat{\xi}_t \leq \tilde{\xi}_t < \xi_t$, any $s \geq t$, and any $(\xi_{t+1}, .., \xi_s)$, let

$$\Delta v_{s,t}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}, .., \xi_s) \equiv v_s(\theta_1, \xi^{t-1}, \tilde{\xi}^+_t, \xi_{t+1}, .., \xi_s) - v_t(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}, .., \xi_s)$$

where $\tilde{\xi}^+_t > \tilde{\xi}_t$ denotes the (upward) shock adjacent to $\tilde{\xi}_t$ while $\tilde{\xi}^-_t < \tilde{\xi}_t$ denotes the (downward) shock adjacent to $\tilde{\xi}_t$. The construction of the transfers in Step 1 then implies that

$$U_t(\theta_1, \xi^{t-1}, \xi_t) = U_t(\theta_1, \xi^{t-1}, \hat{\xi}_t) + \sum_{\tilde{\xi}_t = \hat{\xi}_t}^{\xi_t^-} \left[\Delta v_{t,t}(\theta_1, \xi^{t-1}, \tilde{\xi}_t) q_t(\theta_1, \xi^{t-1}, \tilde{\xi}_t) \right] \tag{43}$$

$$+ \sum_{\tilde{\xi}_t = \hat{\xi}_t}^{\xi_t^-} \left\{ \mathbb{E}_{\xi_{t+1},\ldots,\xi_T} \left[\sum_{s=t+1}^{T} \delta^{s-t} \Delta v_{s,t}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}, .., \xi_s) q_s(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \xi_{t+1}, .., \xi_s) \right] \right\}.$$

Similarly,

$$U_{t+1}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \hat{\xi}_{t+1}^*) - U_{t+1}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \xi_{t+1}) =$$

$$\sum_{\tilde{\xi}_{t+1} = \xi_{t+1}}^{\hat{\xi}_{t+1}^{*-}} \left[\Delta v_{t+1,t+1}(\theta_1, \xi^{t-1}, \tilde{\xi}_t, \tilde{\xi}_{t+1}) q_{t+1}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \tilde{\xi}_{t+1}) \right]$$

$$+ \sum_{\tilde{\xi}_{t+1} = \xi_{t+1}}^{\hat{\xi}_{t+1}^{*-}} \left\{ \mathbb{E}_{\xi_{t+2},\ldots,\xi_T} \left[\sum_{s=t+2}^{T} \delta^{s-t-1} \Delta v_{s,t+1}(\theta_1, \xi^{t-1}, \hat{\xi}_t, \tilde{\xi}_{t+1}, \xi_{t+2}, .., \xi_s) q_s(\theta_1, \xi^{t-1}, \hat{\xi}_t, \tilde{\xi}_{t+1}, \xi_{t+2}, .., \xi_s) \right] \right\}$$

Combining (42) with (43) and using the strong monotonicity property of the quantity schedules, we then have that, when the overlapping condition holds,

$$U_t(\theta_1, \xi^{t-1}, \xi_t) \geq \hat{U}_t(\theta_1, \xi^{t-1}, \xi_t; \theta_1, \xi^{t-1}, \hat{\xi}_t). \tag{44}$$

Similar steps permit us to establish that a buyer who reported truthfully up to period $t-1$ prefers to report truthfully from period t onward rather than reporting *any* shock $\hat{\xi}_t > \xi_t$ in period t and then choosing optimally what to report thereafter.

Finally, that a buyer who reported truthfully in the past finds it optimal to participate in the seller's mechanism at any subsequent date follows directly from the fact that the transfers constructed in Step 1 are such that the type $v_t(\theta_1, \xi^{t-1}, \underline{\underline{\xi}}_t)$ obtains a zero expected continuation payoff. ■

Proof of Proposition 8. Let $\mathcal{F}_t(\theta_t; \theta_1, ..., \theta_{t-1})$ denote the conditional distribution of θ_t given $\theta_1, ..., \theta_{t-1}$. The result is established by induction applying the *probability integral transform theorem* (e.g. Angus, 1994).

Start with $t = 2$. Let ξ_2 be the random variable defined by $\xi_2(\theta_2; \theta_1) \equiv \mathcal{F}_2(\theta_2; \theta_1)$. By the probability integral transform theorem, irrespective of θ_1, ξ_2 is uniformly distributed over $(0, 1)$, which implies that ξ_2 is independent of θ_1. The result for $t = 2$ thus holds by letting $\xi_2 \equiv \mathcal{F}_2(\theta_2; \theta_1)$ and $\theta_2 = v_2(\theta_1, \xi_2) \equiv \mathcal{F}_2^{-1}(\xi_2; \theta_1)$.

Now, by induction, suppose the result holds for all periods $s < t$. I want to show that it holds also for period t. First note that the same arguments used for $t = 2$ imply that there exists a function u_t and a random variable ξ_t such that $\theta_t = u_t(\theta_1, ..., \theta_{t-1}, \xi_t)$ with $\xi_t(\theta_t; \theta_1, ..., \theta_{t-1}) \equiv \mathcal{F}_t(\theta_t; \theta_1, ..., \theta_{t-1})$, $u_t \equiv \mathcal{F}_t^{-1}(\xi_t; \theta_1, ..., \theta_{t-1})$ and ξ_t independent of $\theta_1, ..., \theta_{t-1}$. The fact that each θ_s can be represented as $v_s(\theta_1, \xi^s)$, $s \leq t-1$, then implies that there exists a function $v_t : \mathbb{R}^t \longrightarrow \mathbb{R}$ such that θ_t can be represented as

$$\theta_t = v_t(\theta_1, \xi^t) \equiv \mathcal{F}_t^{-1}(\xi_t; \theta_1, v_2(\theta_1, \xi_2), ..., v_{t-1}(\theta_1, \xi^{t-1}))$$

Furthermore, because for each (θ_1, ξ^{t-1}), $\xi_t \equiv \mathcal{F}_t(\theta_t; v_2(\theta_1, \xi_2), ..., v_{t-1}(\theta_1, \xi^{t-1}))$ is uniformly distributed over $(0,1)$, then ξ_t is independent of (θ_1, ξ^{t-1}). I conclude that (θ_1, ξ^T) are jointly independent. ■

Proof of Proposition 9. *Part (i).* Using the same transformation as in the proof of Proposition 8 adjusted to the fact that θ_t follows a Markov process, we have that θ_t can be represented as

$$\theta_t = v_t(\theta_1, \xi^t) \equiv \mathcal{F}_t^{-1}(\xi_t; v_{t-1}(\theta_1, \xi^{t-1}))$$

with $\xi_t \equiv \mathcal{F}_t(\theta_t; v_{t-1}(\theta_1, \xi^{t-1}))$.

The fact that $\mathcal{F}_t(\theta_t; \theta_{t-1})$ is continuous and strictly increasing in θ_t and strictly decreasing in θ_{t-1} then implies that $\mathcal{F}_t^{-1}$ is strictly increasing in both ξ_t and $v_{t-1}(\theta_1, \xi^{t-1})$. When $t = 2$, this immediately implies that $v_2(\theta_1, \xi_2)$ is increasing in both arguments. The result for any $t > 2$ then follows by induction.

Part (ii). I now want to show that, if the process for θ_t satisfies the conditions in part (i) in the proposition, then any pair of collections $(v_t, \xi_t)_{t=2}^T$ and $(\tilde{v}_t, \tilde{\xi})_{t=2}^T$ that represent this process are such that

$$\frac{\partial v_t(\theta_1, \xi^t)}{\partial \theta_1} = \frac{\partial \tilde{v}_t(\theta_1, \tilde{\xi}^t)}{\partial \theta_1}$$

for any (θ_1, ξ^t) and $(\theta_1, \tilde{\xi}^t)$ such that $\tilde{v}_s(\theta_1, \tilde{\xi}^s) = v_s(\theta_1, \xi^s)$ for any $s \leq t$. Because the process is Markov, there exist functions $(u_t)_{t=2}^T$ and $(\tilde{u}_t)_{t=2}^T$ such that, for any t,

$$v_t(\theta_1, \xi^t) = u_t(v_{t-1}(\theta_1, \xi^{t-1}), \xi_t) \text{ and } \tilde{v}_t(\theta_1, \tilde{\xi}^t) = \tilde{u}_t(\tilde{v}_{t-1}(\theta_1, \tilde{\xi}^{t-1}), \tilde{\xi}_t) \tag{45}$$

with u_t and $\tilde{u}_t$ strictly increasing in ξ_t and $\tilde{\xi}_t$, respectively. Now take any (θ_{t-1}, ξ_t) and let $\tilde{\xi}_t = \tilde{\xi}_t(\theta_{1-1}, \xi_t)$ be implicitly defined by $u_t(\theta_{t-1}, \xi_t) = \tilde{u}_t(\theta_{t-1}, \tilde{\xi}_t)$. I first want to show that

$$\frac{\partial u_t(\theta_{t-1}, \xi_t)}{\partial \theta_{t-1}} = \frac{\partial \tilde{u}_t(\theta_{t-1}, \tilde{\xi}_t(\theta_{t-1}, \xi_t))}{\partial \theta_{t-1}} \tag{46}$$

In fact, because $u_t(\theta_{t-1},\xi_t) = \tilde{u}_t(\theta_{t-1},\tilde{\xi}_t(\theta_{t-1},\xi_t))$ for any (θ_{t-1},ξ_t), we have that

$$\frac{\partial u_t(\theta_{t-1},\xi_t)}{\partial \theta_{t-1}} = \frac{\partial \tilde{u}_t(\theta_{t-1},\tilde{\xi}_t(\theta_{t-1},\xi_t))}{\partial \theta_{t-1}} + \frac{\partial \tilde{u}_t(\theta_{t-1},\tilde{\xi}_t(\theta_{t-1},\xi_t))}{\partial \tilde{\xi}_t}\frac{\partial \tilde{\xi}_t(\theta_{t-1},\xi_t)}{\partial \theta_{t-1}}$$

I now want to show that $\tilde{\xi}_t(\theta_{t-1},\xi_t)$ does not depend on θ_{t-1}. Indeed, suppose this is not the case. Take a pair (θ_{t-1},ξ_t'). Then

$$\Pr(\xi_t \leq \xi_t') = \Pr(\theta_t \leq u_t(\theta_{t-1},\xi_t') \mid \theta_{t-1}) = \Pr(\theta_t \leq \tilde{u}_t(\theta_{t-1},\tilde{\xi}_t(\theta_{t-1},\xi_t')) \mid \theta_{t-1}) = \Pr(\tilde{\xi}_t \leq \tilde{\xi}_t(\theta_{t-1},\xi_t')). \tag{47}$$

Now take a pair (θ_{t-1}',ξ_t') with $\theta_{t-1}' \neq \theta_{t-1}$. Then

$$\Pr(\xi_t \leq \xi_t') = \Pr(\theta_t \leq u_t(\theta_{t-1}',\xi_t') \mid \theta_{t-1}') = \Pr(\theta_t \leq \tilde{u}_t(\theta_{t-1}',\tilde{\xi}_t(\theta_{t-1}',\xi_t')) \mid \theta_{t-1}') = \Pr(\tilde{\xi}_t \leq \tilde{\xi}_t(\theta_{t-1}',\xi_t')). \tag{48}$$

Combining (47) and (48) and using the fact that the random variable $\tilde{\xi}_t$ does not have mass points, we have that necessarily $\tilde{\xi}_t(\theta_{t-1},\xi_t') = \tilde{\xi}_t(\theta_{t-1}',\xi_t')$, that is, the function $\tilde{\xi}_t(\cdot)$ is independent of θ_{t-1} in which case (46) holds.

Now using (46) and (45), we have that

$$\begin{aligned}
\frac{\partial v_t(\theta_1,\xi^t)}{\partial \theta_1} &= \frac{\partial u_t(v_{t-1}(\theta_1,\xi^{t-1}),\xi_t)}{\partial \theta_{t-1}}\frac{\partial u_{t-1}(v_{t-2}(\theta_1,\xi^{t-2}),\xi_{t-1})}{\partial \theta_{t-2}}\cdots\frac{\partial v_2(\theta_1,\xi_2)}{\partial \theta_1} \\
&= \frac{\partial \tilde{u}_t(\tilde{v}_{t-1}(\theta_1,\tilde{\xi}^{t-1}),\tilde{\xi}_t)}{\partial \theta_{t-1}}\frac{\partial \tilde{u}_{t-1}(\tilde{v}_{t-2}(\theta_1,\tilde{\xi}^{t-2}),\tilde{\xi}_{t-1})}{\partial \theta_{t-2}}\cdots\frac{\partial \tilde{v}_2(\theta_1,\tilde{\xi}_2)}{\partial \theta_1} \\
&= \frac{\partial \tilde{v}_t(\theta_1,\tilde{\xi}^t)}{\partial \theta_1}
\end{aligned}$$

for any (θ_1,ξ^t) and (θ_1,ξ^t) such that $\tilde{v}_s(\theta_1,\tilde{\xi}^s) = v_s(\theta_1,\xi^s)$ for any $s \leq t$, which proves the result. ■

Proof of Proposition 10. The fact that the functions $\mathcal{F}_t$ satisfy the properties of Propositions 8 and 9 implies that there exist twice differentiable functions $u_t : \mathbb{R}^2 \longrightarrow \mathbb{R}$ such that, for any $t \geq 2$,

$$v_t(\theta_1,\xi^t) = u_t(v_{t-1}(\theta_1,\xi^{t-1}),\xi_t) = u_t(\theta_{t-1},\xi_t) \tag{49}$$

From Lemma 2 in Eso and Szentes (2007) the condition that $[\partial \mathcal{F}_t(\theta_t;\theta_{t-1})/\partial \theta_{t-1}] / [f_t(\theta_t;\theta_{t-1})]$ is increasing in θ_t is equivalent to the condition that

$$\frac{\partial^2 u_t(\theta_{t-1},\xi_t)}{\partial \theta_{t-1}\partial \xi_t} \leq 0 \ \forall (\theta_1,\xi^t),\ t \geq 2. \tag{50}$$

while the condition that $[\partial\mathcal{F}_t(\theta_t;\theta_{t-1})/\partial\theta_{t-1}]\,/\,[f_t(\theta_t;\theta_{t-1})]$ is increasing in θ_{t-1} is equivalent to the condition that

$$\frac{\partial^2 u_t(\theta_{t-1},\xi_t)}{\partial\theta_{t-1}^2}\frac{\partial u_t(\theta_{t-1},\xi_t)}{\partial\xi_t} \leq \frac{\partial^2 u_t(\theta_{t-1},\xi_t)}{\partial\theta_{t-1}\partial\xi_t}\frac{\partial u_t(\theta_{t-1},\xi_t)}{\partial\theta_{t-1}} \tag{51}$$

Conditions (50) and (51) along with the fact that each u_s is strictly increasing in each argument then imply that

$$\frac{\partial^2 u_t(\theta_{t-1},\xi_t)}{\partial\theta_{t-1}^2} \leq 0 \tag{52}$$

Now, using (49), we have that

$$\begin{aligned}\frac{\partial v_t(\theta_1,\xi^t)}{\partial\theta_1} &= \frac{\partial u_t(v_{t-1}(\theta_1,\xi^{t-1}),\xi_t)}{\partial\theta_{t-1}}\times\frac{\partial u_{t-1}(v_{t-2}(\theta_1,\xi^{t-2}),\xi_{t-1})}{\partial\theta_{t-2}} \\ &\quad\times\cdots\times\frac{\partial v_2(\theta_1,\xi_2)}{\partial\theta_1}\end{aligned} \tag{53}$$

Conditions (52) and (50) along with the fact that both v_s and u_s are strictly increasing in each argument then implies that $v_t(\theta_1,\xi^t)$ is concave in θ_1 and that

$$\frac{\partial^2 v_t(\theta_1,\xi^t)}{\partial\theta_1\partial\xi_s} \leq 0 \ \forall(\theta_1,\xi^t),\ \ s\leq t.$$

That is, the two conditions about the conditional distributions $\mathcal{F}_t(\theta_t;\theta_{t-1})$ stated in the proposition imply that the corresponding functions v_t satisfy the properties of part (i) in Proposition 5.

Next, I want to show that the same conditions imply that the corresponding functions v_s are such that, for any $t\geq 3$, any $s\geq t$ and any $(\theta_1,\xi^t,\xi_{t+1},...,\xi_s)$

$$\frac{\partial^2 v_s(\theta_1,\xi^t,\xi_{t+1},...,\xi_s)}{\partial\theta_1\partial\xi_{t-1}}\frac{\partial v_t(\theta_1,\xi^t)}{\partial\xi_t} \leq \frac{\partial^2 v_s(\theta_1,\xi^t,\xi_{t+1},...,\xi_s)}{\partial\theta_1\partial\xi_t}\frac{\partial v_t(\theta_1,\xi^t)}{\partial\xi_{t-1}}$$

That this is true can be seen directly from (53). ■

References

[1] Angus, John, 1994, "The probability integral transform and related results," *Society for Industrial and Applied Mathematics Review*, 36(4), 653-654.

[2] Baron, D. and D. Besanko, 1984, "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, 1 447-470.

[3] Battaglini, Marco, 2005, "Long-term Contracting with Markovian Consumers," *American Economic Review*, 95(3), 637-658.

[4] Battaglini, Marco and Stephen Coate, 2007, "Pareto Efficient Income Taxation with Stochastic Abilities," *Journal of Public Economics*, forthcoming.

[5] Courty, Pascal and Hao Li, 2000, "Sequential Screening," *Review of Economic Studies*, 67, 697-717.

[6] Eso, Peter and Balazs Szentes, 2007, "Optimal Information Disclosure in Auctions and the Handicap Auction," *Review of Economic Studies*, forthcoming.

[7] Fudenberg Drew and Jean Tirole, 1991, *Game Theory*, MIT Press.

[8] Gibbard, A., 1977, "Manipulation of Voting Schemes," *Econometrica*, 41, 587-601.

[9] Green, J. and J-J. Laffont, 1977, "Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods," *Econometrica*, 45, 427-438.

[10] Myerson, R., 1979, "Incentive Compatibility and the Bargaining Problem," *Econometrica*, 47, 61-73.

[11] Myerson, Roger, 1981, "Optimal Auction Design," *Mathematics Operations Research*, Vol. 6, pp. 58-73.