PROGRESSIVE SCREENING: LONG-TERM CONTRACTING WITH A PRIVATELY KNOWN STOCHASTIC PROCESS

RAPHAEL BOLESLAVKSY† AND MAHER SAID‡

MAY 19, 2011

ABSTRACT: We examine a model of long-term contracting in which the buyer is privately informed about the stochastic process by which her value for a good evolves. In addition, her realized values are also her private information. We characterize the profit-maximizing long-term contract offered by a monopolist in this setting. This optimal contract consists of a menu of deterministic sequences of static contracts. Within each sequence, higher realized values lead to greater quantity provision; however, an increasing proportion of buyer types are excluded over time (eventually leading to inefficient early termination of the relationship). Moreover, the menu choices differ by future generosity, with more costly (up-front) plans guaranteeing greater quantity provision in the future. Thus, the seller screens buyers in the initial period, and then progressively screens additional buyers so as to reduce the information rents paid in future periods.

KEYWORDS: Asymmetric information, Dynamic mechanism design, Long-term contracts, Term life insurance, Sequential screening.

JEL CLASSIFICATION: C73, D82, D86.

†UNIVERSITY OF MIAMI SCHOOL OF BUSINESS, R.BOLESLAVSKY@MIAMI.EDU

‡OLIN BUSINESS SCHOOL, WASHINGTON UNIVERSITY IN ST. LOUIS, SAID@WUSTL.EDU

We thank Atila Abdulkadiroglu, Dirk Bergemann, Alessandro Bonatti, Rahul Deb, Tracy Lewis, Pino Lopomo, Leslie Marx, Hamid Nazerzadeh, Mallesh Pai, Larry Samuelson, Curt Taylor, and Huseyin Yildirim for their many helpful comments and suggestions. This work benefited greatly from the comments of two anonymous referees and the editor, Bruno Biais, and seminar participants at several institutions. The second author also thanks Microsoft Research New England for its generous hospitality during stays in which a significant portion of this work was completed.

1. Introduction

Long-term contracts are a salient feature of a wide variety of economic situations. In many of these settings, the fundamental features of the contractual relationship are not static; rather, these features may change over time. Moreover, while the dynamic nature of the relationship may be acknowledged by all parties involved, the precise nature of the changes may be the private information of only one of the parties; for instance, a seller need not be aware of how her buyers' preferences have evolved, and a downstream retailer need not know the effectiveness of an upstream manufacturer's investments in cost reduction. Thus, the manner in which a contract accounts for these dynamic information asymmetries is crucial, and several recent papers (see, among others, Courty and Li (2000), Battaglini (2005), Pavan (2007), and Pavan, Segal, and Toikka (2009)) have explored this question in a variety of settings.

In the present work, we explore the impact of an *additional* source of private information on the structure and properties of optimal long-term contracts. In particular, we are interested in studying settings in which one party is privately informed not only about the current state of the contracting environment, but also about the manner in which this state evolves. We analyze this issue in the standard setting of the literature, that of an ongoing trading relationship between a monopolist seller and a single consumer. In this relationship, the seller has all of the bargaining power and can credibly commit to the terms of trade for the entire interaction at the outset, while the buyer is privately informed about both her preferences in each period and a parameter of the stochastic process which governs the evolution of her value.

Formally, we characterize the profit-maximizing T-period contract (where T is potentially infinite) for a single seller facing a buyer with single-unit demand whose value evolves according to a stochastic process with a privately known transition probability. We assume that in the initial period, the buyer privately observes the value of a parameter $\lambda \in [0,1]$. In subsequent periods, the buyer privately observes a random shock, and her value in each of these periods is the product of all previously observed shocks. Each shock can take one of two values: it can be either a "good" shock u with probability λ, or a "bad" shock $d < u$ with probability $1-\lambda$. The values u and d are both common knowledge, but the realization of each shock is known only to the buyer. Thus, a buyer with a high value of λ is more likely to experience the good shock in each period, and the distribution of her value at any point in time first-order stochastically dominates the value distribution of a buyer with a lower realization of λ.

At the beginning of the interaction, the seller's goal is to design a long-term contract to maximize her expected profit. We assume that the seller has the ability to fully commit to arbitrary contractual forms, and so the revelation principle allows us to restrict attention, without loss of generality, to the class of direct revelation mechanisms in which the buyer is incentivized to report her private information truthfully.[1] These incentives are required to hold both at the time of initial contracting and in every future period thereafter. Moreover, we assume that the buyer is free to terminate the relationship at any point in time, implying that the buyer's continued participation must be incentivized as well. Thus, the seller's problem in its most general form is to design a

[1]This is in contrast to, for instance, Kennan (2001) and Loginova and Taylor (2008), where the seller's lack of commitment power restricts her to offering only one-period spot contracts.

direct revelation mechanism that maximizes the expected discounted sum of payments made by the buyer, subject to incentive compatibility and individual rationality constraints in each period.[2]

The main result of this paper is that the optimal mechanism can be implemented by a long-term contract with an especially simple structure. In the case where the good is produced in each period at zero cost, the seller (optimally) commits to a finite menu of price plans, each of which presents the buyer with an entry fee and a predetermined sequence of prices for the good in each period.[3] On the basis of λ, the buyer selects a plan and then is free, in each period, to exercise an option to purchase the good at the price specified for that period by her chosen plan.[4] Because prices are fixed by the seller at the beginning of the interaction, this mechanism can be implemented without eliciting any further information from the buyer over the lifetime of the contract—the only information that the seller needs the buyer to reveal is her choice of price plan.

Moreover, each of these price plans begins with a finite "honeymoon" phase. In each period of this phase, the price for the good is the lowest possible value that any buyer could have; after the honeymoon phase ends, the price of the good is multiplied by a factor u in each period. Thus, the buyer is incentivized to purchase the good in each period of the honeymoon phase, regardless of the realizations of her value shocks. After this phase, however, the price rises deterministically, while the buyer's value grows only stochastically—with a sufficiently long time horizon, this serves to (inefficiently) terminate the relationship. As the prices in a plan with a longer honeymoon phase are always (weakly) lower than the prices in a plan with a shorter honeymoon phase, such a plan is more attractive to all buyers, regardless of the probabilities of good and bad shocks. However, the entry fees for the various plans are increasing in the length of their honeymoon phases. So in order to justify paying a larger initial fee, the buyer must anticipate that her future values will be (with high probability) sufficiently high that the lower future prices fully compensate for the initial fee—paying a larger entry fee is justified only if the buyer's probability of good shocks λ is sufficiently high. Thus, the various entry fees and honeymoon phase lengths serve to screen across realizations of λ, while the post-honeymoon-phase rise in prices serves to restrict supply to lower-valued buyers, reducing the rents paid to higher-valued buyers.

As an example of a real-world setting with an information structure of the type described above, consider the market for life insurance. Under the terms of a typical life insurance contract, the insurance provider agrees to pay the buyer's beneficiaries (usually her dependents) a predetermined sum in the case of the policyholder's death, while individuals purchase this insurance in order to provide for their dependents after their passing. Obviously, the probability of death at any point in time is a critical factor in determining the value of life insurance to the individual. Because the individual is privately informed about her physical and mental health, family history, lifestyle, stress levels, and so on, her valuation for a life insurance policy is private. However, she may also have private information about how this value is likely to evolve in the future: for example, she may anticipate receiving a series of promotions (with some probability), each of which will lead to more responsibility, more stress, a more sedentary lifestyle, and more travel, all of which increase

[2]Maximizing social welfare in this setting is trivial, as the buyer's initial-period private information is irrelevant to payoffs post-contracting when the buyer learns her valuations.

[3]These qualititative features of the optimal contract carry over to the case of positive marginal cost; see Section 4.

[4]Thus, the buyer chooses among a set of priority pricing schemes à la Harris and Raviv (1981) or Wilson (1993).

the probability of death (and hence the value for life insurance). Thus, a potential buyer of life insurance will have private information about both her value for an insurance policy as well as the manner in which this value evolves over time.

Translating these features into our model, we interpret $u > 1$ as a negative health shock which increases the buyer's value for life insurance. In the absence of such a shock, the buyer's life insurance remains constant (and so $d = 1$). We may then interpret λ as the probability of a negative health shock in each period. Buyers with higher values of λ are more pessimistic about their future health, and consequently expect higher future values for life insurance.

In this setting, our model predicts that the seller optimally offers the buyer her choice from a set of contracts differentiated by their temporal price profile. In particular, each contract offers a constant price during a honeymoon phase of predetermined length, followed by a steep increase in prices in each period thereafter. This is, in fact, precisely the nature of the contracts we observe in the term life insurance market. Insurance companies typically offer a set of contracts differentiated by length of term, where longer-term contracts have a longer and flatter time profile of premiums over the duration of the term, followed by a steep rise in premiums. Thus, real-world term life insurance contracts match the qualitative features of the optimal contracts predicted by our model. (We elaborate further on this point in Section 4.2.)

We also extend our analysis to the setting in which the seller faces an increasing and convex cost function, while dropping the single-unit demand assumption. Meanwhile, the buyer's valuation shocks are independently drawn from a family of continuous distributions parameterized by λ. We assume that this family is ordered by first-order stochastic dominance, so that larger realizations of λ generate, in expectation, higher valuations. As before, the search for an optimal long-term contract can be restricted to the class of incentive compatible and individually rational direct mechanisms.

While the optimal contract in this more general environment is (as is expected) not quite as simple as that in the single-unit demand case described above, it continues to follow a relatively straightforward form. In the initial period, the buyer is offered her choice from a continuum of contingent price-quantity schedules. For each possible initial-period report of λ, the buyer faces a fixed sequence of menus of price-quantity pairs that screen across her sequence of shocks. As is standard in the nonlinear pricing literature, each of these menus provides greater quantities of the good to buyers that report higher shocks. Moreover, these menus feature more generous quantity provision for buyers with higher reported values of λ, increasing the quantities allocated to buyers that receive a positive amount of the good and excluding a smaller range of realized valuations. However, within a particular sequence of menus, the quantity schedule in a given period is less permissive than the preceding period's menu. The seller progressively screens the buyer and "tightens the screws": the set of period-$(t + 1)$ reports that lead to positive quantities is a subset of the corresponding set of period-t reports. Thus, as in the optimal contract for the single-unit demand case, the seller inefficiently restricts supply in order to extract additional rents, and these restrictions are greater for buyers that report lower values of λ. Meanwhile, the prices corresponding to this optimal long-term contract are similar to those in the discrete model:

within each period's price-quantity menu, the prices are determined entirely by the standard integral payment rule that guarantees incentive compatibility in static settings, depending only on the allocation rule for the period in question. Finally, the entry fees for more permissive menus are higher than those of less permissive menus. Again, in order to justify paying a greater initial entry fee, a buyer must anticipate higher future values—the seller screens across the initial private information by using entry fees and more or less generous sequences of menus, and then progressively screens across values with nonlinear prices in future periods.

As is typically the case in dynamic mechanism design, the primary hurdle we face in solving for the seller's optimal long-term contract is the nature of the incentive compatibility constraints when private information is multidimensional. In particular, incentive compatibility requires that the buyer prefers the truthful reporting of her private information to all potential misreports, including compound multi-stage deviations from truthfulness. This generates a complex and relatively intractable set of constraints that must be satisfied by any optimal contract. The predominant approach in the literature for dealing with this issue is to impose a Markov structure on the evolution of types, which (when combined with a "strong monotonicity" property of the allocation rule) implies that the buyer's optimal strategy following a misreport is to make an additional "corrective" misreport.[5] In our model, however, the buyer's "type" in each period consists of *both* her initial-period private information λ *and* her valuation. Since λ has a persistent impact on the distribution of future values, it is impossible to fully "correct" an initial-period misreport by using additional misreports. Thus, the standard approach is not applicable in the present setting.

We therefore employ an indirect approach to solving for the seller's optimal long-term contract: we solve a relaxed problem that imposes only a restricted set of constraints that are necessarily satisfied by any incentive compatible mechanism, and then show that (in our setting) this restricted set of constraints is, in fact, sufficient for "full" incentive compatibility. More specifically, we impose a set of single-deviation constraints that rule out "one-shot" deviations from truthful reporting. We show that, in the discrete shock setting, the allocation rule that follows from this restricted set of constraints depends only on the buyer's realized value in any period, and not on the particular sequence of shocks generating that value. By pairing this allocation rule with a payment scheme that is also path independent, we are able to decouple the buyers incentives in any one period from those in the next. This guarantees that truthfulness is an optimal continuation strategy for a buyer, regardless of her history of past reports or misreports. This property implies that the restricted class of constraints in our relaxed problem are, in fact, sufficient for "global" incentive compatibility. In the more general setting with continuously distributed valuation shocks, however, the solution to the relaxed version of the seller's problem need not have this property. We therefore present a sufficient condition on the conditional distribution of shocks that guarantees the path-independence of the allocation rule, thereby justifying our approach.

[5] This is the approach of, for instance, Battaglini (2005), Pavan (2007), and Pavan, Segal, and Toikka (2009). Courty and Li (2000) do not require a Markov assumption since they consider only a two-period model, while Kakade, Lobel, and Nazerzadeh (2011) directly assume incentive compatibility after initial-period misreports.

Our paper contributes to the recent, and growing, literature on *optimal* dynamic mechanism design. This literature focuses on the design of profit-maximizing mechanisms in dynamic settings.[6] Much of this recent work focuses on settings where agents arrive and depart dynamically over time while their private information remains fixed, including, for example, the work of Board and Skrzypacz (2010); Gershkov and Moldovanu (2009a); Pai and Vohra (2011); Said (2011); and Vulcano, van Ryzin, and Maglaras (2002). In contrast, our paper joins another strand of the literature where the population of agents is fixed, but their private information changes over time—see, for instance, Baron and Besanko (1984); Battaglini (2005); Besanko (1985); Courty and Li (2000); Deb (2009, 2011); Esö and Szentes (2007); Krähmer and Strausz (2011); and Pavan (2007). The recent work of Pavan, Segal, and Toikka (2009, 2010) presents a broad and general framework subsuming much of the pre-existing literature on dynamic contracting and screening as special cases. Among their many contributions, they provide general necessary conditions for incentive compatibility and optimality in dynamic environments, as well as sufficient conditions in some environments.[7] In contrast to their work, which develops a quite general toolkit for dynamic mechanism design, our present focus is narrower and significantly more applied. In particular, we are concerned with the structure of optimal contracts in a particular setting—one in which the buyer is initially privately informed about the *process* by which her value evolves. Therefore, our work extends the existing dynamic mechanism design literature by considering a novel information structure within a multi-period time horizon, allowing us to examine the dynamics of revenue-enhancing distortions in a new setting.

The most closely related work to our own is the sequential screening model of Courty and Li (2000), who study dynamic price discrimination when consumers are initially uncertain about their future demand but receive new private information before consumption; however, we extend their model in important ways.[8] First, we extend their two-period model to an arbitrary (and potentially infinite) time horizon. This allows us to explore the long-term characteristics of optimal contracts; for instance, the progressive screening, "screw-tightening," and (inefficient) early termination of the relationship by the seller are features of the optimal contract that cannot be intuited from a two-period model. In addition, the longer time horizon necessitates consideration of a richer set of incentive compatibility constraints, as the buyer may choose to misreport her value multiple times in an attempt to take advantage of future contractual terms. This introduces additional technical difficulties in identifying the optimal contract, and we provide a new approach to circumventing such issues. Moreover, since multiple allocations are made over time, we find it more compelling to consider contracts where the buyer pays a price for consumption in each period instead of refund contracts where the buyer pays only an upfront fee and then receives

[6] There is also a parallel literature focusing on the design of *efficient* mechanisms for dynamic settings; see, among others, Athey and Segal (2007a,b); Bergemann and Välimäki (2010); Gershkov and Moldovanu (2009b, 2010a,b); and Kuribko and Lewis (2010). The curious reader is directed to Bergemann and Said (2011), who survey both literatures.

[7] This is complemented by the recent work of Kakade, Lobel, and Nazerzadeh (2011), who provide sufficient conditions for optimality in a particular class of "separable" environments.

[8] Dai, Lewis, and Lopomo (2006) examine similar issues in a procurement setting, while Esö and Szentes (2007) extend the work of Courty and Li (2000) in a different direction and examine a two-period setting in which the seller is able to control the release of new information to multiple buyers. Meanwhile, Miravete (2003) demonstrates empirically the importance of sequential screening considerations in the design of contracts for telephone service.

a refund in the next period. Finally, we extend our results to a setting with infinitely divisible goods and convex costs of production in addition to the standard setting of single-unit demand and constant marginal cost.

This latter feature of our analysis is shared with Battaglini (2005), who also studies long-term contracts when the buyer's type evolves over time. Unlike the present work, his paper considers a setting where the stochastic process governing the buyer's values (a two-state Markov process) is commonly known, but the current state of the process is privately observed by the buyer alone. The optimal contract in that setting converges to the efficient level of supply, sharply contrasting with the inefficiently early termination we find in our setting. This difference stems from the multiplicative structure of the buyer's preferences in our model and the absence of a "highest" value—regardless of the past history of shocks, the buyer's value may increase further, and so the seller continues to distort allocations in order to extract surplus from these higher types. Thus, the cumulative effect of the buyer's new information leads to further restriction of supply over time instead of eventual efficiency. We should point out that this payoff structure also provides a counterpoint to the additive preferences of Krähmer and Strausz (2011), as well as the examples with autoregressive values presented by Pavan, Segal, and Toikka (2009, 2010).

2. Environment

We consider a dynamic setting in which a buyer repeatedly purchases a nondurable good from a single seller. When the buyer in period t pays a price p and receives quantity q of the good, her utility is $v_t q - p$. The buyer's value for the good v_t evolves over time; in particular, we assume that the buyer's value is subject to a series of multiplicative shocks, so that

$$v_t = \alpha_t v_{t-1},$$

where we take $v_0 = 1$ to be exogenously given and commonly known.

In each period $t = 1, \ldots, T$, the buyer privately observes the shocks to her valuation, which are the realizations $\{\alpha_t\}$ of a sequence of independently and identically distributed random variables $\{\tilde{\alpha}_t\}$. In particular, we assume that each shock α_t is drawn from $A := \{u, d\}$, where

$$\Pr(\tilde{\alpha}_t = u) = \lambda \text{ and } \Pr(\tilde{\alpha}_t = d) = 1 - \lambda.$$

We assume that $u > d > 0$, and let $\Delta := u - d$. Thus, the buyer's value evolves according to a recombinant binomial tree process with upward transition probability λ.[9] In each period, the buyer's value experiences either a "good" shock (u) or a "bad" shock (d), and the probability λ of the buyer experiencing the higher shock u in each period is fixed across time.

At the time of contracting (which we take to be period zero), the buyer is privately informed about this transition probability λ. Specifically, the buyer privately observes the realization λ of a random variable $\tilde{\lambda}$, where it is commonly known that $\tilde{\lambda}$ is distributed according to the distribution function F on the interval $\Lambda := [0, 1]$. We denote by f the density of F, and assume throughout that f is strictly positive and differentiable on Λ.

[9] Recombinant binomial tree processes have been frequently used in finance to model asset prices since their introduction by Cox, Ross, and Rubinstein (1979).

In each period $t \geq 1$, the seller can produce q units of the good at a cost $c(q)$.[10] We assume that the relationship between the buyer and the seller is repeated for $T \leq \infty$ periods and discounted with the common discount factor $\delta \in (0, 1]$ (with the additional restriction that $\delta u < 1$ if $T = \infty$). In the initial period (period 0), the seller offers a long-term supply contract to the buyer; the buyer can either accept or reject this offer. If the buyer accepts, sales and consumption occur in periods $t = 1, \ldots, T$ in accordance with the terms of the contract. We normalize the buyer's outside option to 0. As is standard in dynamic models of price discrimination, we assume that the monopolist fully commits to the contract that is offered. However, we will assume that commitment is one-sided: the buyer is free to break off the relationship at any time.

3. The Seller's Problem

The seller in this setting wishes to design and offer the contract that maximizes her expected profits. It is easy to show that in our environment, the Myerson (1986) revelation principle for multistage games holds. Therefore, the search for optimal contracts may be restricted without loss of generality to the class of direct mechanisms where, in each period, the agent is asked to report her new information and, conditional on having reported truthfully in the past, she finds it optimal to report truthfully.

3.1. *Direct Mechanisms*

In particular, a contract in our setting is a sequence of payment rules $\mathbf{p} = \{p_t(r_t, h_t)\}_{t=0}^{T}$ and allocation probabilities $\mathbf{q} = \{q_t(r_t, h_t)\}_{t=1}^{T}$, where r_t is the buyer's report at time t, and h_t is the public history at time t. Note that in such a direct mechanism, $r_0 \in \Lambda$, while $r_t \in A$ for all $t \geq 1$. In addition, h_t can be defined recursively by $h_0 := \emptyset$ and $h_t := \{r_{t-1}, h_{t-1}\}$ for all $t \geq 1$, where r_{t-1} is the agents report in period $t-1$. We denote the set of time t public histories by H_t. Since the agent is free to misreport her private information at any time, her *private* history is $\hat{h}_t := \{\alpha_t, r_{t-1}, \hat{h}_{t-1}\}$, where $\hat{h}_0 := \{\lambda\}$. We denote the set of time t private histories by $\hat{H}_t$; the buyer's strategy, given the seller's mechanism, is then simply a sequence of mappings $\hat{r}_t : \hat{H}_t \to A$ for $t \geq 1$, and $\hat{r}_0 : \Lambda \to \Lambda$.

In addition, we will find it convenient to denote histories using a particular notation. We will denote by α^t the sequence of shocks received by the buyer up to, and including, time t; that is,

$$\alpha^t := (\alpha_t, \alpha_{t-1}, \ldots, \alpha_1).$$

In addition, the notation α^t_{-s} will denote the sequence of shocks up to (and including) period t, but *after* period s, so that

$$\alpha^t_{-s} := (\alpha_t, \alpha_{t-1} \ldots, \alpha_{s+1}).$$

Finally, we will abuse notation somewhat to simplify the exposition and write $v(\alpha^t)$ to denote the value of a buyer who has experienced the sequence of shocks α^t, so that

$$v(\alpha^t) := \prod_{\tau=1}^{t} \alpha_\tau.$$

[10]In Section 4, we will examine the case of single-unit demand for a good produced at a constant marginal cost ($c(q) = cq$), as well as the case of an infinitely divisible good with an increasing and convex cost function ($c(q) = q^2/2$).

A direct mechanism is said to be *incentive compatible* if it induces truthtelling in every period; that is, on the equilibrium path, the agent has no incentive to misreport her new private information. In particular, this implies that the agent prefers revealing her private information truthfully to any potential misreport, followed by optimal reporting in the future (where optimal continuation reporting may involve future misreports). This implies that the set of incentive compatibility constraints in our setting is large and potentially quite complex, and therefore intractable.

In order to avoid this problem, we use an indirect approach to solving for the seller's optimal mechanism. In particular, we consider a restricted set of constraints that are necessarily satisfied by any incentive compatible mechanism, and then show that (in our setting) this restricted set of constraints is, in fact, sufficient for "full" incentive compatibility. More specifically, we require that the buyer prefers reporting her private information truthfully to misreporting in any given period and then reporting truthfully in every future period; that is, we rule out single-period deviations from truthful reporting. The optimal allocation rule that follows from this restricted set of constraints has a "path independence" property (that will be made clear in subsequent sections) that is inherited from the stochastic process governing values. Since there is an additional degree of freedom in choosing payment rules in dynamic mechanisms (relative to their static counterparts), this allocation rule can be paired with a path independent payment rule that guarantees truthtelling as an optimal continuation strategy for a buyer who has misreported in the past, thereby implying the sufficiency of the restricted set of constraints for "global" incentive compatibility.

To state the initial (period 0) single-deviation constraint, let $U_0(\lambda)$ denote the utility of a buyer with initial type λ who always reports her private information truthfully; thus, for all $\lambda \in \Lambda$,

$$U_0(\lambda) := -p_0(\lambda) + \sum_{t=1}^{T} \delta^t \mathbb{E}\left[q_t(\alpha^t, \lambda)v(\alpha^t) - p_t(\alpha^t, \lambda)\middle|\lambda\right]. \tag{1}$$

Similarly, let $\hat{U}_0(\lambda', \lambda)$ denote the expected utility of a buyer with initial type λ who reports some λ', but then truthfully reports *all* future shocks; for all $\lambda, \lambda' \in \Lambda$, we have

$$\hat{U}_0(\lambda', \lambda) := -p_0(\lambda') + \sum_{t=1}^{T} \delta^t \mathbb{E}\left[q_t(\alpha^t, \lambda')v(\alpha^t) - p_t(\alpha^t, \lambda')\middle|\lambda\right]. \tag{2}$$

Thus, the initial period single-deviation constraint requires that

$$U_0(\lambda) \geq \hat{U}_0(\lambda', \lambda) \text{ for all } \lambda, \lambda' \in \Lambda. \tag{IC-0}$$

As with $U_0(\lambda)$, denote by $U_t(\alpha^t, \lambda)$ the expected utility of a buyer in period t whose initial type was λ and whose observed shocks were $\alpha^t \in A^t$, and who has reported truthfully in the past and continues to do so in the present and future. Then

$$\begin{aligned} U_t(\alpha^t, \lambda) &:= q_t(\alpha^t, \lambda)v(\alpha^t) - p_t(\alpha^t, \lambda) \\ &\quad + \sum_{s=t+1}^{T} \delta^{s-t} \mathbb{E}\left[q_s(\alpha^s_{-t}, \alpha^t, \lambda)v(\alpha^s_{-t}, \alpha^t) - p_s(\alpha^s_{-t}, \alpha^t, \lambda)\middle|\alpha^t, \lambda\right]. \end{aligned} \tag{3}$$

Preventing a single deviation in period t requires, for all $(\alpha^t, \lambda) \in A^t \times \Lambda$ and all $\alpha'_t \in A$, that

$$U_t(\alpha^t, \lambda) \geq q_t(\alpha'_t, \alpha^{t-1}, \lambda)v(\alpha^t) - p_t(\alpha'_t, \alpha^{t-1}, \lambda) + \sum_{s=t+1}^{T} \delta^{s-t}\mathbb{E}\left[q_s(\alpha^s_{-t}, \alpha'_t, \alpha^{t-1}, \lambda)v(\alpha^s_{-t}, \alpha^t) - p_s(\alpha^s_{-t}, \alpha'_t, \alpha^{t-1}, \lambda)\middle|\alpha^t, \lambda\right]. \tag{IC-t}$$

In addition, we will say that a direct mechanism is *individually rational* if, in every period and for every history of private signals, it guarantees the buyer's (continued) willingness to participate in the contract by providing expected utility greater than her outside option. These individual rationality constraints may be summarized by the following:

$$U_0(\lambda) \geq 0 \text{ for all } \lambda \in \Lambda, \text{ and} \tag{IR-0}$$

$$U_t(\alpha^t, \lambda) \geq 0 \text{ for all } (\alpha^t, \lambda) \in A^t \times \Lambda \text{ and all } t = 1, \ldots, T. \tag{IR-t}$$

The seller's profit from any feasible contract is the difference between total surplus and the buyer's utility. Thus, when the buyer is of initial type λ, the seller's expected profit is

$$\begin{aligned}\Pi(\lambda) &:= p_0(\lambda) + \sum_{t=1}^{T} \delta^t \mathbb{E}\left[p_t(\alpha^t, \lambda) - c(q_t(\alpha^t, \lambda))\middle|\lambda\right] \\ &= -U_0(\lambda) + \sum_{t=1}^{T} \delta^t \mathbb{E}\left[q_t(\alpha^t, \lambda)v(\alpha^t) - c(q_t(\alpha^t, \lambda))\middle|\lambda\right] \end{aligned}\tag{4}$$

The seller's optimal contract maximizes profits, subject to the constraints that the consumer receives at least her reservation utility and that the consumer has no incentive to misreport her type. Thus, any optimal contract must also solve the *relaxed* problem that imposes the individual rationality constraints and the restricted set of single-deviation incentive compatibility constraints:

$$\max_{\{\mathbf{p},\mathbf{q}\}} \left\{ \int_0^1 \Pi(\lambda)\, \mathrm{d}F(\lambda) \right\} \quad \text{subject to (IC-0), (IR-0), (IC-}t\text{), and (IR-}t\text{) for all } t = 1, \ldots, T. \tag{$\mathcal{P}$}$$

3.2. *Simplifying the Seller's Relaxed Problem*

We approach the seller's optimal contracting problem by simplifying the single-deviation and participation constraints in the relaxed problem ($\mathcal{P}$). Recall that in a standard static contracting setting, the combination of quasilinearity and a single-crossing condition imply that incentive constraints are equivalent to (i) the monotonicity of the allocation rule; and (ii) the determination of a buyer's utility (up to a constant) by that allocation rule alone. In addition, the single-crossing property and individual rationality constraint pin down this constant. We begin by showing that a similar technique may be applied in our setting to the period-t constraints. Note, however, that the buyer is forward looking, and her utility depends upon her expectations (given her information in period t) over future shocks. Naturally, this implies that the "localized" period-t constraints will involve the expected discounted value of current and future allocations, which we denote by

$$\bar{q}_t(\alpha^t, \lambda) := q_t(\alpha^t, \lambda)v(\alpha^{t-1}) + \sum_{s=t+1}^{T} \delta^{s-t}\mathbb{E}\left[q_s(\alpha^s_{-t}, \alpha^t, \lambda)v(\alpha^s_{-t}, \alpha^{t-1})\middle|\alpha^t, \lambda\right]. \tag{5}$$

Thus, we have the following "standard" result (whose proof may be found in appendix):

LEMMA 3.1. *The period-t incentive compatibility and individual rationality constraints* (IC-t) *and* (IR-t), *where* $t = 1, \dots, T$, *are satisfied if, and only if, for all* $\alpha^{t-1} \in A^{t-1}$ *and all* $\lambda \in \Lambda$,

$$U_t(u, \alpha^{t-1}, \lambda) - U_t(d, \alpha^{t-1}, \lambda) \geq \bar{q}_t(d, \alpha^{t-1}, \lambda)\Delta; \tag{IC$'$-t}$$

$$\bar{q}_t(u, \alpha^{t-1}, \lambda) \geq \bar{q}_t(d, \alpha^{t-1}, \lambda); \text{ and} \tag{MON-t}$$

$$U_t(d, \alpha^{t-1}, \lambda) \geq 0. \tag{IR$'$-t}$$

Notice that at the time of initial contracting (in contrast to any period $t \geq 1$), the buyer's private information does not directly affect her payoffs. Rather, the realization of λ only provides the buyer a (noisy) signal about the evolution of her future preferences. Therefore, the buyer has preferences over the entire sequence of allocations, and so we cannot appeal to a single-crossing condition to simplify the initial-period constraints. However, using an envelope argument (detailed in the appendix), we can show that period-zero single-deviation constraint necessarily implies that the buyer's interim (in the initial period) expected utility depends only upon the the expectation of future payoff gradients; in particular, this observation—in conjunction with the period-t single-deviation constraints (IC$'$-t)—will then allow us to reformulate the seller's relaxed problem ($\mathcal{P}$) into one involving only allocation rules (and not the payment rules).

LEMMA 3.2. *If the period-zero incentive compatibility constraint (IC-0) is satisfied, then the derivative* $U_0'(\lambda)$ *of the buyer's period-zero expected utility is given by*

$$U_0'(\lambda) = \sum_{t=1}^{T} \delta^t \mathbb{E}\left[U_t(u, \alpha^{t-1}, \lambda) - U_t(d, \alpha^{t-1}, \lambda)\Big|\lambda\right]. \tag{IC$'$-0}$$

Note that since allocation probabilities are non-negative, if condition (IC$'$-t) is satisfied, $U_0'(\lambda)$ is also non-negative. Therefore, U_0 must be increasing in any solution to the seller's problem, and we can replace the period-zero individual rationality constraint (IR-0) with the requirement that

$$U_0(0) \geq 0. \tag{IR$'$-0}$$

We should point out that (IC$'$-0) is a necessary implication of period-zero incentive compatibility, but that it is *not* sufficient. Although this expression allows us to simplify the optimization problem by eliminating transfers from the seller's objective function, those transfers remain a part of the problem's constraints.

Finally, we may return to the seller's problem. Since (IC$'$-0) must hold in any incentive compatible mechanism, we may use the standard integration by parts methodology to reformulate the optimization problem ($\mathcal{P}$) as

$$\max_{\{\mathbf{q},\mathbf{p}\}} \left\{ \begin{aligned} &-U_0(0) + \sum_{t=1}^{T} \delta^t \int_\Lambda \mathbb{E}\left[q_t(\alpha^t, \lambda)v(\alpha^t) - c(q_t(\alpha^t, \lambda))\big|\lambda\right] \,\mathrm{d}F(\lambda) \\ &\quad - \sum_{t=1}^{T} \delta^t \int_\Lambda \frac{1-F(\lambda)}{f(\lambda)} \mathbb{E}\left[U_t(u, \alpha^{t-1}, \lambda) - U_t(d, \alpha^{t-1}, \lambda)\Big|\lambda\right] \,\mathrm{d}F(\lambda) \end{aligned} \right\}$$

subject to (IC-0), (IC$'$-t), (MON-t), (IR$'$-0), and (IR$'$-t) for all $t = 1, \dots, T$.

Clearly, any solution to this problem must have $U_0(0) = 0$, as this is merely an additive constant that is bounded by the constraint (IR′-0)—as is standard, providing additional surplus to the lowest type only reduces the seller's profit without helping provide incentives for truthtelling. In addition, it is clear that solving the problem requires minimizing $U_t(u, \alpha^{t-1}, \lambda) - U_t(d, \alpha^{t-1}, \lambda)$ for each $\alpha^{t-1} \in A^{t-1}$ and $\lambda \in \Lambda$. However, the constraints (IC′-t) provide a lower bound on this difference, and so these constraints must bind. In particular, this implies that we may incorporate these constraints into the seller's objective function, rewriting it as

$$\sum_{t=1}^{T} \delta^t \int_\Lambda \mathbb{E}\left[q_t(\alpha^t, \lambda) v(\alpha^t) - \bar{q}_t(d, \alpha^{t-1}, \lambda)\Delta \frac{1-F(\lambda)}{f(\lambda)} - c(q_t(\alpha^t, \lambda)) \middle| \lambda \right] \mathrm{d}F(\lambda).$$

Finally, note that

$$\begin{aligned}\sum_{t=1}^{T} \delta^t \mathbb{E}\left[\bar{q}_t(d, \alpha^{t-1}, \lambda) \middle| \lambda\right] &= \sum_{t=1}^{T} \delta^t \mathbb{E}\Bigg[q_t(d, \alpha^{t-1}, \lambda) v(\alpha^{t-1}) \\ &\qquad + \sum_{s=t+1}^{T} \delta^{s-t} \mathbb{E}\left[q_s(\alpha^s_{-t}, d, \alpha^{t-1}, \lambda) v(\alpha^s_{-t}, \alpha^{t-1}) \middle| \alpha^{t-1}, \lambda \right] \Bigg| \lambda \Bigg] \\ &= \sum_{t=1}^{T} \sum_{s=t}^{T} \delta^s \mathbb{E}\left[q_s(\alpha^s_{-t}, d, \alpha^{t-1}, \lambda) v(\alpha^s_{-t}, \alpha^{t-1}) \middle| \lambda \right] \\ &= \sum_{t=1}^{T} \sum_{s=1}^{t} \delta^t \mathbb{E}\left[q_t(\alpha^t_{-s}, d, \alpha^{s-1}, \lambda) v(\alpha^t_{-s}, \alpha^{s-1}) \middle| \lambda \right], \end{aligned} \tag{6}$$

where the final equality follows from interchanging the order of summations. Substituting this expression back into the seller's profit-maximization problem yields the following relaxed problem:

$$\max_{\{\mathbf{q},\mathbf{p}\}} \left\{ \sum_{t=1}^{T} \delta^t \int_\Lambda \mathbb{E}\left[\begin{array}{l} q_t(\alpha^t, \lambda) v(\alpha^t) - c(q_t(\alpha^t, \lambda)) \\ \quad - \displaystyle\sum_{s=1}^{t} q_t(\alpha^t_{-s}, d, \alpha^{s-1}, \lambda) v(\alpha^t_{-s}, \alpha^{s-1}) \Delta \frac{1-F(\lambda)}{f(\lambda)} \end{array} \middle| \lambda \right] \mathrm{d}F(\lambda) \right\} \tag{$\mathcal{P}'$}$$

subject to (IC-0), (MON-t), and (IR′-t) for all $t = 1, \ldots, T$.

3.3. *Virtual Values*

Before proceeding to the solution of the seller's problem, it is helpful to interpret the seller's objective function in ($\mathcal{P}'$), especially by way of comparison with a standard nonlinear pricing setting (as in, for instance, Mussa and Rosen (1978)). For each t and every $\lambda \in \Lambda$, we can rewrite the integrand in the objective function as

$$\begin{aligned} &\sum_{\alpha^t \in A^t} \Pr(\alpha^t | \lambda) \left(v(\alpha^t) q_t(\alpha^t, \lambda) - c(q_t(\alpha^t, \lambda)) \right) \\ &\quad - \sum_{\alpha^t \in A^t} \Delta \frac{1-F(\lambda)}{f(\lambda)} \sum_{s=1}^{t} \Pr(\alpha^t_{-s}, \alpha^{s-1} | \lambda) v(\alpha^t_{-s}, \alpha^{s-1}) \mathbb{1}_d(\alpha_s) q_t(\alpha_t, \lambda) \\ &\qquad = \sum_{\alpha^t \in A^t} \Pr(\alpha^t | \lambda) \left(v(\alpha^t) \left[1 - \sum_{s=1}^{t} \mathbb{1}_d(\alpha_s) \frac{\Delta / d}{1-\lambda} \frac{1-F(\lambda)}{f(\lambda)} \right] q_t(\alpha_t, \lambda) - c(q_t(\alpha^t, \lambda)) \right), \end{aligned}$$

where $\mathbb{1}_d(\alpha_s)$ is the indicator function for the event $\{\alpha_s = d\}$. Thus, the seller is essentially maximizing, in the Myersonian tradition, virtual surplus, where the buyer's virtual value is

$$\varphi(\alpha^t, \lambda) := v(\alpha^t)\left[1 - \sum_{s=1}^{t} \mathbb{1}_d(\alpha_s)\frac{\Delta/d}{1-\lambda}\frac{1-F(\lambda)}{f(\lambda)}\right] = v(\alpha^t) - v(\alpha^t)\sum_{s=1}^{t} \mathbb{1}_d(\alpha_s)\frac{\Delta/d}{1-\lambda}\frac{1-F(\lambda)}{f(\lambda)}. \quad (7)$$

As in the static mechanism design setting, the first term in this expression is the buyer's contribution to the social surplus, while the second term represents the information rents that must be "paid" to the buyer in order to induce truthful revelation of her private information. The inverse hazard rate $(1 - F(\lambda))/f(\lambda)$ appears since any information rents paid to a buyer with initial type λ must also be paid to buyers with higher initial types. Finally, as is well-established in the dynamic mechanism design literature, distortions in the optimal contract are determined by the sensitivity of future values to the buyer's initial private information. In our setting, the responsiveness of the buyer's period-t value to her private information at the time of contracting is given by $v(\alpha^t)\sum_{s=1}^{t} \mathbb{1}_d(\alpha_s)\frac{\Delta/d}{1-\lambda}$; more specifically, this expression is equal to the derivative $\partial v_t/\partial\lambda$ of the period-t value v_t with respect to λ. Since our environment is a hybrid of continuous and discrete type spaces, this property is most easily demonstrated by using the "independent shock approach" introduced in Esö and Szentes (2007) and further developed in Pavan (2007).

With this in mind, define $\{\xi_t\}_{t=1}^{T}$ to be a sequence of independent (across t and, crucially, also from λ) draws from a uniform distribution on $[0, 1]$, and let

$$\tilde{v}_t(\xi^t, \lambda) := \begin{cases} d\tilde{v}_{t-1}(\xi^{t-1}, \lambda) & \text{if } \xi_t < 1-\lambda, \\ u\tilde{v}_{t-1}(\xi^{t-1}, \lambda) & \text{if } \xi_t \geq 1-\lambda, \end{cases}$$

where ξ^s denotes the s-tuple $(\xi_s, \xi_{s-1}, \ldots, \xi_1)$ and we define $\tilde{v}_0(\lambda) := 1$ for all λ. Thus, we can identify all $\xi_t < 1 - \lambda$ with the shock $\alpha_t = d$, and similarly, we can identify all $\xi_t \geq 1 - \lambda$ with the shock $\alpha_t = u$. This implies that we may identify a buyer with private type (α^t, λ) with the "average buyer" whose private type is (ξ^t, λ) with ξ_τ corresponding to α_τ for all $\tau = 1, \ldots, t$.

Since we can write

$$\tilde{v}_t(\xi^t, \lambda) = (d + \Delta H_{1-\lambda}(\xi_t))\,\tilde{v}_{t-1}(\xi^{t-1}, \lambda),$$

where $H_z(\cdot)$ is the Heaviside step function shifted by $z \in \mathbb{R}$, it is straightforward to see that

$$\frac{\partial\tilde{v}_t(\xi^t, \lambda)}{\partial\lambda} = (d + \Delta H_{1-\lambda}(\xi_t))\,\frac{\partial\tilde{v}_{t-1}(\xi^{t-1}, \lambda)}{\partial\lambda} + \Delta\delta_{1-\lambda}(\xi_t)\tilde{v}_{t-1}(\xi^{t-1}, \lambda),$$

where $\delta_z(\cdot)$ denotes the Dirac delta centered at $z \in \mathbb{R}$. Then

$$\begin{aligned}\mathbb{E}\left[\frac{\partial\tilde{v}_t(\xi^t, \lambda)}{\partial\lambda}\middle|\xi_t < 1-\lambda\right] &= d\frac{\partial\tilde{v}_{t-1}(\xi^{t-1}, \lambda)}{\partial\lambda} + \Delta\frac{\int_0^{1-\lambda} \mathrm{d}H_{1-\lambda}(\xi_t)}{\int_0^{1-\lambda} \mathrm{d}\xi_t}\tilde{v}_{t-1}(\xi^{t-1}, \lambda) \\ &= d\frac{\partial\tilde{v}_{t-1}(\xi^{t-1}, \lambda)}{\partial\lambda} + \frac{\Delta}{1-\lambda}\tilde{v}_{t-1}(\xi^{t-1}, \lambda),\end{aligned}$$

while

$$\mathbb{E}\left[\frac{\partial \tilde{v}_t(\xi^t,\lambda)}{\partial\lambda}\middle|\xi_t \geq 1-\lambda\right] = u\frac{\partial \tilde{v}_{t-1}(\xi^{t-1},\lambda)}{\partial\lambda} + \Delta\frac{\int_{1-\lambda}^{1} \mathrm{d}H_{1-\lambda}(\xi_t)}{\int_{1-\lambda}^{1} \mathrm{d}\xi_t}\tilde{v}_{t-1}(\xi^{t-1},\lambda)$$
$$= u\frac{\partial \tilde{v}_{t-1}(\xi^{t-1},\lambda)}{\partial\lambda}.$$

Proceeding recursively (and noting that $\partial\tilde{v}_0(\lambda)/\partial\lambda = 0$), this implies that

$$v(\alpha^t)\sum_{\tau=1}^{t}\mathbb{1}_d(\alpha_\tau)\frac{\Delta/d}{1-\lambda} = \mathbb{E}\left[\frac{\partial\tilde{v}_t(\xi^t,\lambda)}{\partial\lambda}\middle| \begin{array}{l}\xi_s < 1-\lambda \text{ if } \alpha_s = d,\\ \xi_s \geq 1-\lambda \text{ if } \alpha_s = u.\end{array}\right].$$

Therefore, the virtual value of a buyer with type (α^t,λ) is given by

$$\varphi(\alpha^t,\lambda) = v(\alpha^t)\left[1-\sum_{s=1}^{t}\mathbb{1}_d(\alpha_s)\frac{\Delta/d}{1-\lambda}\frac{1-F(\lambda)}{f(\lambda)}\right].$$

Note that the fact that each shock enters the buyer's value multiplicatively implies that the additional distortions generated by the dependence of α_t on λ have the additive structure above. This is in contrast to, for example, the nature of distortions in Besanko (1985), the Markovian setting of Battaglini (2005), the autoregressive values examples presented by Pavan, Segal, and Toikka (2009, 2010), and the procurement model of Krähmer and Strausz (2011).

Before moving on, we make the following additional assumption:

Assumption A. *The distribution of initial-period values F is such that*

$$\frac{1-F(\lambda)}{(1-\lambda)f(\lambda)}$$

is (weakly) decreasing in λ.

This regularity assumption is a sufficient condition for the buyer's virtual value $\varphi(\alpha^t,\lambda)$ to be nondecreasing in λ for all $\alpha^t \in A^t$ and all t, and will allow us to avoid concerns about ironing that arise when virtual values are non-monotone. This assumption is satisfied by a large variety of distributions F on $[0,1]$. For instance, the uniform distribution on $[0,1]$, as well as any power distribution $F(\lambda) = \lambda^x$, where $x \geq 1$, satisfies this assumption. Similarly, the beta and Kumaraswamy distributions satisfy Assumption A whenever their shape parameters are $a \geq 1$ and $b > 0$. We should note that this assumption is strictly stronger than the standard regularity assumption that F is log-concave; however, similar regularity conditions are imposed in essentially all dynamic mechanism design settings.[11]

[11]The utility of such conditions was first noted by Baron and Besanko (1984) and Besanko (1985), and an analogous condition was imposed by Matthews and Moore (1987) in a multi-dimensional screening setting. An equivalent condition on the conditional distribution of future types on initial period private information is found in, for instance, the sequential screening model of Courty and Li (2000), the information disclosure setting of Esö and Szentes (2007), and the Krähmer and Strausz (2011) model of pre-project planning in procurement. Pavan, Segal, and Toikka (2009) also make additional monotonicity conditions on primitives in order to avoid ironing concerns. A notable exception to this trend is found in Kakade, Lobel, and Nazerzadeh (2011); however, their model imposes "separability" conditions that rule out the dependence of buyer values on initial-period private information found in the present work.

4. The Optimal Long-Term Contract

We now solve for the optimal dynamic mechanism in this setting for two important cases. We first consider the case in which the buyer has single-unit demand and the good is produced at a constant marginal cost. In this setting, the optimal contract has an especially simple structure: the seller commits to a menu of deterministic price sequences. We then turn to the case of an increasing and convex cost function, and relax the single-unit demand assumption. In this case, the optimal long-term contract has a more sophisticated, but qualitatively similar, structure: the seller commits to a menu of deterministic sequences of price-quantity schedules. In both cases, the initial menu screens across the buyer's initial-period private information, while the sequence of prices (or schedules) progressively screen across the buyer's realized valuations.

4.1. *Single-Unit Demand and Constant Marginal Cost*

Suppose first that the seller produces the good in each period at a constant marginal cost, so that $c(q) = cq$ for some constant $c \geq 0$. In addition, suppose that the buyer has single-unit demand. In this case, the seller's optimization problem becomes

$$\max_{\{\mathbf{q},\mathbf{p}\}} \left\{ \sum_{t=1}^{T} \delta^t \int_{\Lambda} \mathbb{E}\left[(\varphi(\alpha^t, \lambda) - c) q_t(\alpha^t, \lambda) \middle| \lambda \right] \, \mathrm{d}F(\lambda) \right\}$$

subject to (IC-0), (MON-t), and (IR$'$-t).

Notice that for all t, the seller's objective function is linear in $q_t(\alpha^t, \lambda)$ for all $\alpha^t \in A^t$ and $\lambda \in \Lambda$. Therefore (temporarily ignoring the constraints (IC-0), (MON-t), and (IR$'$-t)), the seller sets $q_t(\alpha^t, \lambda) = 1$ if, and only if, $\varphi(\alpha^t, \lambda) \geq c$, and otherwise sets $q_t(\alpha^t, \lambda) = 0$.

Consider a history (α^t, λ) where $\sum_{s=1}^{t} \mathbb{1}_d(\alpha_s) = k$, and note that the condition $\varphi(\alpha^t, \lambda) \geq c$ may be rewritten as

$$u^{t-k} d^k \left(1 - \frac{\Delta/d}{1-\lambda} \frac{1-F(\lambda)}{f(\lambda)} k \right) \geq c.$$

So for all $\lambda < 1$, $\frac{\Delta/d}{1-\lambda}\frac{1-F(\lambda)}{f(\lambda)} > 0$, implying that the left-hand side of this inequality is decreasing in k. Therefore, for every $t = 1, \ldots, T$ and every initial-period type λ, the buyer is allocated an object as long as she has experienced sufficiently few downward shocks d.[12] Formally, we define

$$k_t(\lambda) := \max \left\{ k \in \mathbb{Z}_+ : u^{t-k} d^k \left(1 - \frac{\Delta/d}{1-\lambda} \frac{1-F(\lambda)}{f(\lambda)} k \right) \geq c \right\}, \tag{8}$$

where we let $k_t(\lambda) := 0$ if the set being maximized over is empty. The cutoff $k_t(\lambda)$ is finite for all $\lambda < 1$. To see this, note that as long as F has a derivative of *any* order that is non-zero when $\lambda = 1$, l'Hôpital's rule implies that

$$\lim_{\lambda \to 1} \frac{\Delta/d}{1-\lambda} \frac{1-F(\lambda)}{f(\lambda)} = \gamma \text{ for some constant } \gamma > 0; \tag{9}$$

[12]Note that, for a buyer with initial-period type $\lambda = 1$, $\varphi(\alpha^t, \lambda) = v(\alpha^t)$ for all $\alpha^t \in A^t$; thus, such a buyer's allocation experiences no distortions away from the efficient allocation. We do not focus on this, however, as this is a zero-probability type.

thus, for any t, a buyer who has experienced more than $1/\gamma$ downward shocks d will have a *negative* virtual value. Clearly, our regularity assumption on the distribution F implies that the cutoff $k_t(\lambda)$ is non-decreasing in λ for all t.

Finally, let us define

$$q_t^*(\alpha^t, \lambda) := \begin{cases} 1 & \text{if } \sum_{s=1}^t \mathbb{1}_d(\alpha_s) \leq k_t(\lambda), \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

Since the term $\sum_{s=1}^t \mathbb{1}_d(\alpha_s)$ simply counts the number of realized d shocks in a given history of signals α^t, it is trivial to see that $q_t^*(u, \alpha^{t-1}, \lambda) \geq q_t^*(d, \alpha^{t-1}, \lambda)$ for all $\alpha^{t-1} \in A^{t-1}$ and $\lambda \in \Lambda$. Moreover, this also implies that $q_s^*(\alpha_{-t}^s, u, \alpha^{t-1}, \lambda) \geq q_s^*(\alpha_{-t}^s, d, \alpha^{t-1}, \lambda)$ for all $s > t$ and all $\alpha_{-t}^s \in A^{s-t}$ and $\lambda \in \Lambda$. Since this holds for every realization of α_{-t}^s, it must also hold when taking expectations (given λ), and therefore condition (MON-t) is satisfied. This fact, combined with the fact that the constraints (IC′-t) bind, implies that the complete set of period-t (for $t \geq 1$) single-deviation constraints (IC-t) are satisfied.[13]

Of course, these single-deviation constraints are only a (necessary) subset of the full set of incentive constraints that must be satisfied. In particular, the constraints (IC-t) guarantee only that the buyer prefers reporting her type truthfully in period $t \geq 1$ to a *single* deviation from truthfulness; this property is not, in general, sufficient to guarantee that the buyer does not wish to misreport her type *multiple* times over the course of the relationship. However, the allocation rule in Equation (10) depends only on the number of downward shocks d the buyer has experienced, but *not* the order in which they were received—q_t^* is path independent. This observation suggests that combining the optimal allocation rule with a path-independent payment rule may lead to "full" incentive compatibility.

The payment scheme we propose is essentially a sequence of prices determined by the standard (static) Myersonian payment rule applied to the entire range of possible values in each period, and not just those that are possible given a particular history of reports. Thus, the "price" of the good in each period $t \geq 1$ is simply the lowest possible reported period-t value for which the buyer still receives the good:

$$p_t^*(\alpha^t, \lambda) := \begin{cases} u^{t-\min\{t, k_t(\lambda)\}} d^{\min\{t, k_t(\lambda)\}} & \text{if } \sum_{s=1}^t \mathbb{1}_d(\alpha_s) \leq k_t(\lambda), \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

Having fixed a payment scheme for all future periods, the period-zero "entry fee" is easily pinned down. Using the definition of $U_0(\lambda)$ in Equation (1), we may write

$$p_0(\lambda) = \sum_{t=1}^T \delta^t \mathbb{E}\left[q_t(\alpha^t, \lambda) v(\alpha^t) - p_t(\alpha^t, \lambda) \middle| \lambda\right] - U_0(\lambda).$$

[13]We will show shortly (see Theorem 1) that the remaining constraints (IC-0) and (IR′-t) in the seller's relaxed problem ($\mathcal{P}'$) are also satisfied, as is the complete set of incentive compatibility constraints.

Since $U_0(\lambda) = \int_0^\lambda U_0'(\mu)\,\mathrm{d}\mu$, we can use Lemma 3.2 (combined with the fact that the constraints (IC′-t) and (IR′-0) bind) to show that the initial payment must be

$$p_0^*(\lambda) := \sum_{t=1}^{T} \delta^t \mathbb{E}\left[q_t^*(\alpha^t,\lambda)(v(\alpha^t) - p_t^*(\alpha^t,\lambda))\middle|\lambda\right] - \int_0^\lambda \sum_{t=1}^{T} \delta^t \mathbb{E}\left[\bar{q}_t^*(d,\alpha^{t-1},\mu)\Delta\middle|\mu\right]\,\mathrm{d}\mu. \tag{12}$$

Note that this contract $(\mathbf{q}^*, \mathbf{p}^*)$ guarantees that $p_t^*(\alpha^t,\lambda) \leq q_t^*(\alpha^t,\lambda)v(\alpha^t)$ for all $(\alpha^t,\lambda) \in A^t \times \Lambda$, and so the buyer's expected flow payoff (when truthful) in each period is always non-negative. Therefore, the individual rationality constraints (IR′-t) are all satisfied.

One natural way to think about the allocation and payment rules above is to consider the corresponding indirect mechanism: the seller can implement the contract described above by giving the buyer a choice among several "plans" differentiated by their initial up-front cost and future sequence of prices. In each period after the initial choice of plan, the seller does not elicit any further information from the buyer, but instead simply presents her with a deterministic sequence of prices. Since the buyer's behavior after the initial period does not affect future prices, she can simply make the myopically optimal choice of purchasing the good in period t if the price is lower than her value.

This elimination of dynamic incentives is precisely the feature of the proposed contract that guarantees satisfaction of the "full" set of incentive compatibility constraints: the contract induces truthful reporting by the buyer even after histories in which she previously misreported her private information (be it λ or α_t for some t). This is because a period-t misreport (for $t \geq 1$) has one of two effects: over-reporting the number of d shocks leads to the exclusion of the buyer in situations where truthful reporting may have led to a profitable allocation, while under-reporting the number of d shocks leads to allocations at prices greater than the buyer's value. As neither of these two outcomes affects future prices or values, the buyer has no ability to manipulate the mechanism in future periods, and so there is neither a static nor dynamic incentive for misreporting one's value.

Thus, it only remains to verify that the proposed solution satisfies the initial-period single-deviation constraint (IC-0). As previously noted, the "localized" version of the constraint derived in Lemma 3.2 is generally only a necessary, but not sufficient, condition for period-zero incentive compatibility. However, since it guarantees the monotonicity of the allocation in λ, Assumption A is sufficient for incentive compatibility in the initial period. The theorem below (whose proof is found in the appendix) demonstrates this fact.

Theorem 1. *Suppose that the distribution F satisfies Assumption A. Then the contract $(\mathbf{q}^*, \mathbf{p}^*)$, where $\mathbf{q}^*$ denotes the quantity schedules from Equation (10) and $\mathbf{p}^*$ denotes the payment rules from Equations* (11) *and* (12)*, is an optimal contract that solves the seller's problem ($\mathcal{P}'$).*

So as to fully appreciate the optimal mechanism proposed above, it is helpful to consider the special case where the good is produced at zero cost in each period. In this case, the condition $\varphi(\alpha^t,\lambda) \geq c$ is equivalent to the requirement that

$$\frac{\Delta/d}{1-\lambda}\frac{1-F(\lambda)}{f(\lambda)}\sum_{s=1}^{t}\mathbb{1}_d(\alpha_s) \leq 1.$$

Thus, the optimal allocation rule is time independent, and simply sets an upper bound $\bar{k}(\lambda)$ on the number of downward shocks d permitted over the course of the relationship for every period-zero report λ. Moreover, given Assumption A, the optimal contract partitions the set of initial-period types into a set of intervals $\Lambda_n := [\lambda_{n-1}, \lambda_n)$ such that $\bar{k}(\lambda) = n$ for all $\lambda \in \Lambda_n$. Each of these intervals corresponds to a "plan" of future price paths offered by the seller.

Within each plan, the path of prices is straightforward, with the price changing at a predetermined rate in each period—in the plan designated for a buyer with $\lambda \in \Lambda_n$, the price grows by a factor d in each of the first n periods, and then grows by a factor u in *every* period thereafter. (Technically, the price need not actually "grow": if $u < 1$ or $d < 1$, prices will be decreasing.) This initial period of slower growth in prices is essentially a "honeymoon" phase, after which the slope of the price path increases. Thus, the set of plans offered by the seller vary by the length of their honeymoon phases, with longer honeymoon phases demanding higher entry fees in the initial period. Indeed, in order to justify paying a larger initial fee, the buyer must anticipate that her future values will be (with high probability) sufficiently high that the lower future prices fully compensate her for the upfront cost—paying a larger initial fee for a future price discount is justified only if the buyer's probability λ of experiencing upward shocks u is sufficiently high.

Additionally, it is important to note that the length of the honeymoon phase in each plan is finite, as is the number of plans offered. (This finiteness follows from the observation in Equation (9).) Thus, the seller never finds it optimal to continue serving a buyer after they have experienced a fixed *finite* number of downward shocks, regardless of the number of upward shocks already experienced. Furthermore, note that $\bar{k}(\lambda)$ is independent of the length of the time horizon T (as well as the discount factor δ). This implies that early (inefficient) termination of the contract will occur with probability arbitrarily close to 1 given a sufficiently long time horizon T. Indeed, the law of large numbers implies that, for all $\lambda < 1$, the probability of the buyer experiencing more than $\bar{k}(\lambda)$ downward d shocks in the first $n < T$ periods approaches 1 as n grows large. Once this occurs, the buyer will make no additional payments, and will never again receive the good. Thus, the seller commits to early termination of the relationship so as to increase her revenue.

When the cost of producing the good is strictly positive, then the optimal allocation rule q_t^* need not be time independent, nor does the seller necessarily offer a finite number of plans. In particular, $k_t(\lambda) \geq k_{t+1}(\lambda)$ when $u < 1$, and $k_t(\lambda) \leq k_{t+1}(\lambda)$ when $u > 1$. In this latter case, a buyer with a virtual value that is positive but less than the marginal cost c will be excluded, but if her value recovers with sufficiently many u shocks, she may be allocated the object again. However, note that since $\varphi(\alpha^t, \lambda) \leq 0$ whenever $\sum_{s=1}^{t} \mathbb{1}_d(\alpha_s) > \bar{k}(\lambda)$, we must have $k_t(\lambda) \leq \bar{k}(\lambda)$ for all t, where $\bar{k}(\lambda)$ is the upper bound from the costless production case discussed above—once the buyer's virtual value becomes negative, it remains negative and the buyer is excluded in all future periods. Thus, the "price" of the good will eventually grow deterministically at the higher rate u, while the buyer's value will only probabilistically grow at that rate—as time proceeds, the seller progressively screens the buyer by restricting supply when she receives a downward shock d so as to extract additional rents from the buyer when she receives the higher u shocks. With a sufficiently long time horizon T, this rent extraction leads to the eventual exclusion of all buyers.

4.2. *Term Life Insurance*

Recall the life insurance example discussed in the introduction, where a buyer's value for life insurance at any point in time varies, as it is influenced by the probability of death. Moreover, recall that this probability of death may be influenced by privately observed factors whose evolution is also the potential policyholder's private information. So letting λ denote the probability of a negative health shock in each period, and letting $u > 1 =: d$, we may interpret our model as one in which the buyer's health (and hence value for insurance) remains constant unless she receives a negative health shock. (Buyers with higher values of λ therefore have more pessimistic expectations about their future health, and are likely to have higher future values for life insurance.)

In such a setting, our model predicts several features of the optimal long-term contract. First, the optimal contract can be implemented without eliciting information from the buyer over the lifetime of the relationship. Second, the seller can implement this optimal contract by committing to a menu of deterministic price plans, each element of which is differentiated by the length of its honeymoon phase. Moreover, plans with longer honeymoon phases involve a higher upfront fee but (weakly) lower per-period prices.

Table 1 demonstrates the annualized premiums and cumulative payments, sampled over the course of 65 years, for each of the Select Term 10 Year, 20 Year, and 30 Year renewable term life insurance policies offered by the State Farm Life Insurance Company.[14] The initial premium for each plan is guaranteed for the length of the term (ten, twenty, and thirty years, respectively), during which no health information is elicited. After this term, the policyholder is free to renew the policy *without* providing a statement of health or additional proof of insurability to State Farm; however, the premiums "increase significantly" and "continue to increase annually and are adjustable but will not exceed the maximum premiums listed in the policy." As with the term life insurance contracts studied by Hendel and Lizzeri (2003), there is one-sided commitment: the insurer is legally bound to the terms of the contract while the policyholder may costlessly terminate it at any time.

	Select Term 10 Year		Select Term 20 Year		Select Term 30 Year	
Year	Premium	Cumulative	Premium	Cumulative	Premium	Cumulative
5	335	1,675	350	1,750	525	2,625
10	335	3,350	350	3,500	525	5,250
15	3,365	17,385	350	5,250	525	7,875
20	4,850	38,905	350	7,000	525	10,500
30	12,815	124,695	12,815	92,790	525	15,750
35	22,100	214,495	22,100	182,590	22,100	105,550
50	96,440	990,070	96,440	958,165	96,440	881,125
60	264,680	2,763,960	264,680	2,732,055	264,680	2,655,015
65	385,295	4,440,145	385,295	4,408,240	385,295	4,331,200

TABLE 1. State Farm Term Life Insurance Policy Premiums

[14]Premiums are the "Elite Preferred Non-Tobacco" rates, calculated for $500,000 of coverage for a 30-year-old male in the state of Missouri. This data can be obtained by visiting http://www.statefarm.com/insurance/quote/lrq.asp. Similarly structured term life insurance policies are described by, for instance, Hendel and Lizzeri (2003).

Thus, State Farm (among other insurers) offers buyers a choice of period of time in which premiums remain constant, followed by an increasing premium thereafter. It should be clear that the premiums during the term of the 20- and 30-year plans may be reduced to the initial premiums of the 10-year plan without changing the cumulative expenditure by charging an upfront entry fee equal to the discounted value of the payment difference.[15] Moreover, these longer-term policies cost more (over the course of their term) than the shorter-term 10-year policy. Thus, these term life insurance policies appears to closely match the optimal long-term contract prescribed by our model; while we do not want to suggest that State Farm acts as though it were a monopolist, the natural interpretation is that the insurer is using the different terms to screen across buyers with differing future health profiles, while raising the premiums over time to progressively screen and filter out the consumers with the lowest willingness to pay for insurance.

4.3. *Convex Costs*

The results presented and discussed above are not limited to the unit-demand setting above; a similar contractual structure arises in a setting in which the seller faces an increasing convex cost function and we relax the assumption of single-unit demand. To see this, consider the case where the seller can produce q units of the good in each period at a cost of $c(q) = q^2/2$.[16] Then the seller's problem ($\mathcal{P}'$) may be written as

$$\max_{\{\mathbf{q},\mathbf{p}\}} \left\{ \sum_{t=1}^{T} \delta^t \int_\Lambda \mathbb{E}\left[\varphi(\alpha^t,\lambda) q_t(\alpha^t,\lambda) - \frac{q_t^2(\alpha^t,\lambda)}{2} \middle| \lambda \right] \, \mathrm{d}F(\lambda) \right\}$$
$$\text{subject to (IC-0), (MON-}t\text{), and (IR}'\text{-}t\text{).}$$

Pointwise maximization (for each (α^t,λ) tuple) of the integrand while ignoring (for now) the constraints yields the following solution:

$$q_t^*(\alpha^t,\lambda) := \max\left\{ v(\alpha^t)\left(1 - \sum_{s=1}^{t} \mathbb{1}_d(\alpha_s)\frac{\Delta/d}{1-\lambda}\frac{1-F(\lambda)}{f(\lambda)}\right), 0 \right\}. \tag{13}$$

Notice that this allocation rule distorts the buyer's quantity away from the first-best (efficient) allocation by a factor that depends on the number of downward shocks d that the buyer reports. Thus, a report of d in period t affects the buyer's allocation in two ways: first, it leads to a decrease in her reported value (relative to the value that would have been inferred from a report of u), thereby decreasing the quantity she would have been allocated in a complete-information setting; and second, it leads to an increase in the distortion away from the efficient allocation. Moreover, both of these effects carry through to the allocation in all future periods. Therefore, for every $t = 1, \ldots, T$,

$$q_t^*(u,\alpha^{t-1},\lambda) \geq q_t^*(d,\alpha^{t-1},\lambda)$$

[15]The 20-year plan above is equivalent to a 20-year policy with an annual premium of \$335 and an upfront fee of $\sum_{\tau=1}^{20}(350-335)/(1+r)^\tau$, where r is the interest rate. Similarly, the 30-year plan is equivalent to a 30-year policy with an annual premium of \$335 and an upfront fee of $\sum_{\tau=1}^{30}(525-335)/(1+r)^\tau$.

[16]Assuming a quadratic cost function, as in Battaglini (2005), Pavan (2007), and much of the static contracting literature, yields significant gains in tractability and ease of exposition without significant loss of generality.

for all $\alpha^{t-1} \in A^{t-1}$ and $\lambda \in \Lambda$, and, for all $s > t$,

$$q_s^*(\alpha_{-t}^s, u, \alpha^{t-1}, \lambda) \geq q_s^*(\alpha_{-t}^s, d, \alpha^{t-1}, \lambda)$$

for all $\alpha_{-t}^s \in A^{s-t}$ and $\lambda \in \Lambda$. Since this latter inequality holds for every realization of α_{-t}^s, it also holds in expectation (conditional on λ), and therefore the constraint (MON-t) is satisfied. Since the constraints (IC′-t) also bind, this implies that the complete set of period-t (with $t \geq 1$) single-deviation incentive constraints are satisfied.

Again, we must note that the satisfaction of these constraints need not, in general, guarantee that the buyer prefers truthful reporting of her type to (potentially complicated) compound deviations. However, as was the case in Section 4.1, the allocation rule defined in Equation (13) is, essentially, a function of λ and the buyer's reported period-t *value* alone—for each $\lambda \in \Lambda$ and all t, $q_t^*(\alpha^t, \lambda) = q_t^*(\hat{\alpha}^t, \lambda)$ for any $\alpha^t, \hat{\alpha}^t \in A^t$ such that $v(\alpha^t) = v(\hat{\alpha}^t)$. Therefore, we make use of a path-independent payment rule in order to incentivize the buyer to treat her reporting decision in any period $t \geq 1$ as a single-period (static) problem.

To this end, we make use of the standard (static) nonlinear pricing rule á la Mussa and Rosen (1978); however, instead of applying this pricing rule to the set of possible values conditional on the reported history α^{t-1} (that is, over the set $\{uv(\alpha^{t-1}, dv(\alpha^{t-1})\}$), we apply it to the *entire* set of possible period-t values $\{u^t, u^{t-1}d, \ldots, ud^{t-1}, d^t\}$. Thus, letting

$$m(\alpha^t) := \sum_{s=1}^{t} \mathbb{1}_d(\alpha_s),$$

we define, for all $t = 1, \ldots, T$ and all $(\alpha^t, \lambda) \in A^t \times \Lambda$,

$$p_t^*(\alpha^t, \lambda) := q_t^*(\alpha^t, \lambda)v(\alpha^t) - \sum_{j=m(\alpha^t)+1}^{t} q_t^*(\underbrace{d, \ldots, d}_{j}, \underbrace{u, \ldots, u}_{t-j})\Delta u^{t-j} d^{j-1}. \tag{14}$$

Note that, with the payments defined above, the buyer's flow payoff in each period (assuming truthful reporting of α^t) is

$$q_t^*(\alpha^t, \lambda)v(\alpha^t) - p_t^*(\alpha^t, \lambda) = \sum_{j=m(\alpha^t)+1}^{t} q_t^*(\underbrace{u, \ldots, u}_{t-j}, \underbrace{d, \ldots, d}_{j})\Delta u^{t-j} d^{j-1} \geq 0.$$

Therefore, the individual rationality constraints (IR′-t) are satisfied for all $t \geq 1$. Moreover, the initial-period payment $p_0^*(\lambda)$ is uniquely determined by combining the definition of $U_0(\lambda)$ in Equation (1) with the envelope condition from Lemma 3.2:

$$p_0^*(\lambda) := \sum_{t=1}^{T} \delta^t \mathbb{E}\left[q_t^*(\alpha^t, \lambda)(v(\alpha^t) - p_t^*(\alpha^t, \lambda))\middle|\lambda\right] - \int_0^\lambda \sum_{t=1}^{T} \delta^t \mathbb{E}\left[\bar{q}_t^*(d, \alpha^{t-1}, \mu)\Delta\middle|\mu\right] \mathrm{d}\mu. \tag{15}$$

Again, it is helpful to interpret the direct mechanism above by considering its indirect counterpart. In the period zero, the seller offers the buyer her choice from a menu of options

$$\{p_0^*(\lambda), \{q_t^*(\cdot, \lambda), p_t^*(\cdot, \lambda)\}_{t=1}^{T}\}_{\lambda \in \Lambda},$$

where each period-zero menu choice consists of an entry fee and a *predetermined* sequence of price-quantity schedules. Then, in each period $t \geq 1$, the buyer is free to choose any of the $t+1$ price-quantity pairs on the period-t schedule that correspond to the $t+1$ possible values in period t. Crucially, her choice in any period $t \geq 1$ does *not* alter the prices or quantities available to her in any future periods. This implies that, given any initial-period report of λ, the buyer's decision problem in each period $t \geq 1$ is decoupled from her decision problem in any other period $t' \geq 1$. Her choice of price-quantity pair then (myopically) maximizes her flow utility in that period.

Notice, however, that since $q_t^*(\cdot, \lambda)$ is decreasing in the reported number of downward d shocks for all t and all λ, it is increasing in the buyer's *value*. Standard results from static mechanism design then imply that the period-t menu is incentive compatible (in the static sense), regardless of the buyer's initial-period report, and so the buyer will choose the price-quantity pair that corresponds to her true value.[17] Thus, for any initial-period report λ, the contract described in Equations (13), (14), and (15) is "fully" incentive compatible: the buyer has no incentive to ever misreport her shocks, even when multiple deviations are permitted.

Of course, this observation does *not* imply that the initial-period single-deviation constraint (IC-0) is satisfied—recall that the envelope condition derived in Lemma 3.2 is only a necessary implication of period-zero incentive compatibility. However, Assumption A implies that the quantity schedules are (weakly) increasing in λ for all t and all possible reports $\alpha^t \in A^t$. The following theorem (with proof in the appendix) shows that this property is, in fact, sufficient to guarantee that the buyer reports truthfully in the initial period, and therefore the incentive compatibility of the proposed contract.

THEOREM 2. *Suppose that the distribution F satisfies Assumption A. Then the contract $(\mathbf{q}^*, \mathbf{p}^*)$, where $\mathbf{q}^*$ denotes the quantity schedules from Equation (13) and $\mathbf{p}^*$ denotes the payment rules from Equations* (14) *and* (15)*, is an optimal contract that solves the seller's problem ($\mathcal{P}'$).*

In this setting, the regularity assumption on the distribution F implies that q_t^* is monotone increasing in λ; therefore, the seller's menu is infinite. However, as in the indivisible goods case discussed in Section 4.1, the optimal contract allows only a fixed finite number of reported d shocks before permanently excluding the buyer, and this upper bound depends only on the buyer's reported initial-period type λ. Thus, each additional d shock reported by the buyer not only decreases the quantity she is allocated, but it also brings her closer to contract termination. Since such shocks occur with strictly positive probability whenever $\lambda < 1$, this outcome is essentially unavoidable given a long enough time horizon T.

5. CONTINUOUS SHOCKS

In order to investigate the robustness of the results presented in the previous section, we now present a more general formulation of the model. Instead of a setting with discrete shocks, we

[17] Note that this does *not* imply that the period-t menu is the optimal menu for (statically) screening across the buyer's potential period-t values in a setting where λ is commonly known.

examine a setting where the buyer's valuation shocks in each period are drawn from a continuous distribution. For concreteness, we focus on the case where the seller faces an increasing and convex cost function.

5.1. *Environment*

As before, we consider a dynamic setting in which the buyer repeatedly purchases a nondurable good from a single seller. The buyer's utility in period t from consuming q units bought at a price p is given by $v_t q - p$. As before, the buyer's marginal value for an additional unit v_t is subject to a series of multiplicative shocks, so that $v_t = \alpha_t v_{t-1}$, where we take $v_0 = 1$ to be exogenously given and commonly known.

In each period $t = 1, \ldots, T$, the buyer privately observes the shocks to her valuation, which are the realizations $\{\alpha_t\}$ of a sequence of random variables $\{\tilde{\alpha}_t\}$, independently and identically drawn from the conditional distribution $G(\cdot|\lambda)$ on the interval $A := [\underline{\alpha}, \bar{\alpha}]$ with $0 \leq \underline{\alpha} < \bar{\alpha} \leq \infty$. We denote by $g(\cdot|\lambda)$ the conditional density of $G(\cdot|\lambda)$, and assume that $g(\cdot|\lambda)$ is strictly positive and differentiable on A for all $\lambda \in \Lambda$. Moreover, we assume that the family $\{G(\cdot|\lambda)\}_\lambda$ is ordered in terms of first-order stochastic dominance; that is, $G(\cdot|\lambda)$ first-order stochastically dominates $G(\cdot|\lambda')$ whenever $\lambda > \lambda'$. In addition, we assume, for technical reasons, that all partial derivatives of G exist and are bounded. Note that this, combined with the previous assumption, implies that $\partial G(\alpha|\lambda)/\partial\lambda \leq 0$.

At the time of contracting (period zero), the buyer is privately informed about the parameter λ of the distribution that generates the sequence of shocks $\{\alpha_t\}$. Specifically, the buyer privately observes the realization λ of a random variable $\tilde{\lambda}$, where it is commonly known that $\tilde{\lambda}$ is distributed according to the distribution function F on an interval $\Lambda := [\underline{\lambda}, \bar{\lambda}]$ with $0 \leq \underline{\lambda} < \bar{\lambda} \leq \infty$. We assume that f, the density of F, is strictly positive and differentiable on Λ.

We assume that the seller can produce q units of the good in each period at a cost of $c(q) := q^2/2$. The buyer-seller relationship is repeated for $T \leq \infty$ periods and payoffs are discounted with a common discount factor $\delta \in (0, 1]$ (where we impose the additional restriction that $\delta\bar{\alpha} < 1$ if $T = \infty$). The timing remains identical to that described in Section 2: in period zero, the seller offers a long-term contract to the buyer; the buyer can either accept or reject this offer. If accepted, the contract is executed in periods $t = 1, \ldots, T$. The buyer's outside option is normalized to 0. We continue to assume that the monopolist enjoys full commitment to the contract she offers, but that the buyer is free to break off the relationship at any time.

5.2. *The Seller's Problem*

As in Section 3, the seller wishes to design and offer the contract that maximizes her expected profits. The Myerson (1986) revelation principle for multistage games continues to hold, and so the search for optimal contracts is again greatly simplified (without loss of generality) by the consideration of direct mechanisms where the buyer is incentivized to report truthfully conditional on having reported truthfully in the past.

Thus, a contract in our setting is a sequence of payment rules $\mathbf{p} = \{p_t(r_t, h_t)\}_{t=0}^T$ and allocation probabilities $\mathbf{q} = \{q_t(r_t, h_t)\}_{t=1}^T$, where r_t is the buyer's report at time t, and h_t is the public history at time t. Note that in such a direct mechanism, $r_0 \in \Lambda$, while $r_t \in A$ for all $t \geq 1$. In addition, h_t can be defined recursively by $h_0 := \emptyset$ and $h_t := \{r_{t-1}, h_{t-1}\}$ for all $t \geq 1$, where r_{t-1} is the agents report in period $t-1$. We denote the set of time t public histories by H_t. Since the agent is free to misreport her private information at any time, her *private* history is $\hat{h}_t := \{\alpha_t, r_{t-1}, \hat{h}_{t-1}\}$, where $\hat{h}_0 := \{\lambda\}$. We denote the set of time t private histories by $\hat{H}_t$; the buyer's strategy, given the seller's mechanism, is then simply a sequence of mappings $\hat{r}_t : \hat{H}_t \to A$ for $t \geq 1$, and $\hat{r}_0 : \Lambda \to \Lambda$.

A mechanism is incentive compatible if, having reported truthfully in previous periods, the agent has no incentive to misreport her new private information. As in Section 3, this implies that the seller's optimal mechanism must prevent the buyer from making multi-stage deviations from truthtelling. So as to avoid considering this large class of potentially complex deviations, we consider a relaxed version of the seller's problem that considers only one-shot deviations from truthfulness. We will then present a sufficient condition that guarantees the optimal allocation rule for this relaxed problem is path independent. Our main result in this section shows that combining this allocation rule with a path-independent payment scheme yields an optimal mechanism that guarantees the (full) incentive compatibility of the proposed mechanism.

With this in mind, let $U_0(\lambda)$ denote the utility of a buyer with initial type λ who always reports her private information truthfully; thus, for all $\lambda \in \Lambda$, we have

$$U_0(\lambda) := -p_0(\lambda) + \sum_{t=1}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda)v(\alpha^t) - p_t(\alpha^t, \lambda)\right) \, \mathrm{d}W^t(\alpha^t|\lambda), \tag{16}$$

where $\mathrm{d}W^t(\alpha^t|\lambda) = \prod_{\tau=1}^{t} \mathrm{d}G(\alpha_\tau|\lambda)$. Similarly, let $\hat{U}_0(\lambda', \lambda)$ denote the expected utility of a buyer with initial type λ who reports some λ', but then truthfully reports all future shocks:

$$\hat{U}_0(\lambda', \lambda) := -p_0(\lambda') + \sum_{t=1}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda')v(\alpha^t) - p_t(\alpha^t, \lambda')\right) \, \mathrm{d}W^t(\alpha^t|\lambda). \tag{17}$$

Thus, the initial-period single-deviation constraint requires that

$$U_0(\lambda) \geq \hat{U}_0(\lambda', \lambda) \text{ for all } \lambda, \lambda' \in \Lambda. \tag{IC-0}$$

Similarly, denote by $U_t(\alpha_t, \alpha^{t-1}, \lambda)$ the expected utility of a buyer who observes the period-t shock α_t after truthfully reporting (α^{t-1}, λ), and then continues to report truthfully. Then

$$\begin{aligned} U_t(\alpha_t, \alpha^{t-1}, \lambda) &:= q_t(\alpha_t, \alpha^{t-1}, \lambda)v(\alpha^t) - p_t(\alpha_t, \alpha^{t-1}, \lambda) \\ &+ \sum_{s=t+1}^{T} \delta^{s-t} \int_{A^{s-t}} \Big(q_s(\alpha^s_{-t}, \alpha_t, \alpha^{t-1}, \lambda)v(\alpha^s_{-t}, \alpha^t) - p_s(\alpha^s_{-t}, \alpha_t, \alpha^{t-1}, \lambda)\Big) \, \mathrm{d}W^{s-t}(\alpha^s_{-t}|\lambda). \end{aligned} \tag{18}$$

The period-t single-deviation constraint requires, for all truthful histories $(\alpha^t, \lambda) \in A^t \times \Lambda$, the superiority of continued truthfulness to a single misreport $\alpha'_t \in A$:

$$\begin{aligned} U_t(\alpha_t, \alpha^{t-1}, \lambda) &\geq q_t(\alpha'_t, \alpha^{t-1}, \lambda)v(\alpha^t) - p_t(\alpha'_t, \alpha^{t-1}, \lambda) \\ &+ \sum_{s=t+1}^{T} \delta^{s-t} \int_{A^{s-t}} \Big(q_s(\alpha^s_{-t}, \alpha'_t, \alpha^{t-1}, \lambda)v(\alpha^s_{-t}, \alpha^t) - p_s(\alpha^s_{-t}, \alpha'_t, \alpha^{t-1}, \lambda)\Big) \, \mathrm{d}W^{s-t}(\alpha^s_{-t}|\lambda). \end{aligned} \tag{IC-t}$$

In addition, a mechanism is individually rational if, in every period, continued truthtelling guarantees the buyer's continued willingness to participate in the contract. These individual rationality constraints may be summarized by the following:

$$U_0(\lambda) \geq 0 \text{ for all } \lambda \in \Lambda, \text{ and} \tag{IR-0}$$

$$U_t(\alpha^t, \lambda) \geq 0 \text{ for all } (\alpha^t, \lambda) \in A^t \times \Lambda \text{ and all } t = 1, \ldots, T. \tag{IR-t}$$

The seller's profit from any feasible contract is the difference between total surplus and the buyer's utility. Thus, when the buyer is of initial type λ, the seller's expected profit is

$$\Pi(\lambda) := -U_0(\lambda) + \sum_{t=1}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda)v(\alpha^t) - c(q_t(\alpha^t, \lambda))\right) \, \mathrm{d}W^t(\alpha^t|\lambda). \tag{19}$$

The seller's optimal contract must then maximize profits, subject to the constraints that the consumer receives at least her reservation utility and that the consumer have no incentive to misreport her type. Therefore, any optimal long-term contract must solve the relaxed problem that imposes only the participation and single-deviation truthtelling constraints:

$$\max_{\{\mathbf{p},\mathbf{q}\}} \left\{ \int_{\Lambda} \Pi(\lambda) \, \mathrm{d}F(\lambda) \right\} \quad \text{subject to (IC-0), (IC-}t\text{), (IR-0), and (IR-}t\text{).} \tag{$\mathcal{Q}$}$$

5.3. *Simplifying the Seller's Relaxed Problem*

We approach the optimal contracting problem by incorporating the period-t (where $t \geq 1$) single-deviation and participation constraints into the objective function. As the buyer's utility is quasilinear and satisfies the standard single-crossing conditions, the single-deviation constraints are equivalent to (i) monotonicity of the buyer's expected discounted allocation; and (ii) the determination of the buyer's utility entirely by that expected allocation, up to a constant. In addition, the single-crossing property pins down this constant by the individual rationality constraint. Since the period-t buyer is forward-looking, her utility depends upon her expectations (in period t) over future shocks in periods $s \geq t$. Naturally, this implies that the "localized" version of the period-t constraints will involve the expected discounted allocations for $s \geq t$, which we denote by

$$\begin{aligned}\bar{q}_s(\alpha_t, \alpha^{t-1}, \lambda) := {} & q_t(\alpha_t, \alpha^{t-1}, \lambda)v(\alpha^{t-1}) \\ & + \sum_{s=t+1}^{T} \delta^{s-t} \int_{A^{s-t}} q_s(\alpha^s_{-t}, \alpha_t, \alpha^{t-1}, \lambda)v(\alpha^s_{-t}, \alpha^{t-1}) \, \mathrm{d}W^{s-t}(\alpha^{s-t}|\lambda).\end{aligned} \tag{20}$$

The standard machinery yields the following result, whose proof may be found in the appendix:

Lemma 5.1. *The period-t single-deviation and individual rationality constraints* (IC-t) *and* (IR-t) *are satisfied if, and only if, for all $t = 1, \ldots, T$, and all $(\alpha^{t-1}, \lambda) \in A^{t-1} \times \Lambda$,*

$$\frac{\partial}{\partial \alpha_t} U_t(\alpha_t, \alpha^{t-1}, \lambda) = \bar{q}_t(\alpha_t, \alpha^{t-1}, \lambda) \text{ for all } \alpha_t \in A; \tag{IC$'$-t}$$

$$\bar{q}_t(\alpha_t, \alpha^{t-1}, \lambda) \text{ is nondecreasing in } \alpha_t; \text{ and} \tag{MON-t}$$

$$U_t(\underline{\alpha}, \alpha^{t-1}, \lambda) \geq 0. \tag{IR$'$-t}$$

Note that the buyer's private information at the time of initial contracting does not directly affect her payoffs. Rather, λ only provides information about the path of future preferences over the entire sequence of allocations. Therefore, the standard single-crossing condition does not apply, and we must resort instead to an envelope argument in order to simplify the seller's problem. Using the following result (with proof in the appendix), we are able to remove the payment rules from the objective function in the seller's relaxed problem ($\mathcal{Q}$). As in the case of discrete shocks discussed in Section 3, this envelope condition is a necessary implication of period-zero incentive compatibility, but it is *not* in general sufficient.

LEMMA 5.2. *Suppose that the single-deviation constraints* (IC-0) *and* (IC-t) *are satisfied for all t. Then the derivative $U_0'(\lambda)$ of the buyer's period-zero expected utility is given by*

$$U_0'(\lambda) = -\sum_{t=1}^{T} \delta^t \int_{A^t} \bar{q}_t(\alpha^t, \lambda) \frac{\partial G(\alpha_t|\lambda)/\partial\lambda}{g(\alpha_t|\lambda)} \, \mathrm{d}W^t(\alpha^t|\lambda). \tag{IC$'$-0}$$

Moreover, if the single-deviation constraints are satisfied, then the period-zero incentive rationality constraint (IR-0) is equivalent to the requirement that

$$U_0(\underline{\lambda}) \geq 0. \tag{IR$'$-0}$$

Finally, we return to the seller's problem. Since (IC$'$-0) must hold in any incentive compatible mechanism, integration by parts implies that the relaxed problem ($\mathcal{Q}$) may be rewritten as

$$\max_{\{\mathbf{q},\mathbf{p}\}} \left\{ \begin{aligned} &-U_0(\underline{\lambda}) + \int_\Lambda \sum_{t=1}^{T} \delta^t \int_{A^t} \bar{q}_t(\alpha^t, \lambda) \frac{\partial G(\alpha_t|\lambda)/\partial\lambda}{g(\alpha_t|\lambda)} \frac{1-F(\lambda)}{f(\lambda)} \, \mathrm{d}W^t(\alpha^t|\lambda) f(\lambda) \, \mathrm{d}\lambda \\ &\quad + \int_\Lambda \sum_{t=1}^{T} \delta^t \int_{A^t} \big(q_t(\alpha^t,\lambda) v(\alpha^t) - c(q_t(\alpha^t,\lambda)) \big) \, \mathrm{d}W^t(\alpha^t|\lambda) f(\lambda) \, \mathrm{d}\lambda \end{aligned} \right\}$$

subject to (IC-0), (MON-t), (IR$'$-0), and (IR$'$-t).

Since $U_0(\underline{\lambda})$ is an additive constant in the objective function above, it must be the case that the individual rationality constraint (IR$'$-0) binds. Moreover, note that

$$\begin{aligned} &\sum_{t=1}^{T} \delta^t \int_{A^t} \bar{q}_t(\alpha^t, \lambda) \frac{\partial G(\alpha_t|\lambda)/\partial\lambda}{g(\alpha_t|\lambda)} \, \mathrm{d}W^t(\alpha^t|\lambda) \\ &\qquad = \sum_{t=1}^{T} \delta^t \int_{A^t} q_t(\alpha^t, \lambda) \sum_{s=1}^{t} v(\alpha^t_{-s}, \alpha^{s-1}) \frac{\partial G(\alpha_s|\lambda)/\partial\lambda}{g(\alpha_s|\lambda)} \, \mathrm{d}W^t(\alpha^t|\lambda). \end{aligned} \tag{21}$$

Thus, the relaxed version of the seller's problem becomes

$$\max_{\{\mathbf{q},\mathbf{p}\}} \left\{ \begin{aligned} &\sum_{t=1}^{T} \delta^t \iint_{\Lambda \times A^t} \Bigg(q_t(\alpha^t, \lambda) \left[v(\alpha^t) + \sum_{s=1}^{t} v(\alpha^t_{-s}, \alpha^{s-1}) \frac{\partial G(\alpha_s|\lambda)/\partial\lambda}{g(\alpha_s|\lambda)} \frac{1-F(\lambda)}{f(\lambda)} \right] \\ &\qquad\qquad - c(q_t(\alpha^t, \lambda)) \Bigg) \, \mathrm{d}W^t(\alpha^t|\lambda) \, \mathrm{d}F(\lambda) \end{aligned} \right\} \tag{$\mathcal{Q}'$}$$

subject to (IC-0), (MON-t), and (IR$'$-t).

5.4. *The Optimal Contract*

As in Section 3 (and as is standard in optimal mechanism design more generally), the seller here is essentially maximizing virtual surplus, where the buyer's virtual value in period $t = 1, \dots, T$ is

$$\varphi(\alpha^t, \lambda) := v(\alpha^t) + \sum_{s=1}^{t} v(\alpha^t_{-s}, \alpha^{s-1}) \frac{\partial G(\alpha_s|\lambda)/\partial\lambda}{g(\alpha_s|\lambda)} \frac{1 - F(\lambda)}{f(\lambda)} \tag{22}$$

The first term in each of these expressions is the buyer's contribution to the social surplus, while the second term represents the information rents that must be left to the buyer in order to induce truthful revelation of her private information. The inverse hazard rate $(1 - F(\lambda))/f(\lambda)$ appears since any information rents paid to a buyer with initial type λ must also be paid to buyers with higher initial types. Meanwhile, the additional

$$\frac{\partial G(\alpha|\lambda)/\partial\lambda}{g(\alpha|\lambda)}$$

terms are the "informativeness measures" of Baron and Besanko (1984), which reflect the informativeness of the initial-period private information λ on future shocks α_s, where we sum over all $s \leq t$ to account for the different shocks through which λ influences $v(\alpha^t)$.[18] Note that first-order stochastic dominance implies that $\partial G(\alpha|\lambda)/\partial\lambda \leq 0$, and so these cumulative information rents are not paid *by* the buyer, but rather *to* her; therefore, $\varphi(\alpha^t, \lambda) \leq v(\alpha^t)$ for all $(\alpha^t, \lambda) \in A^t \times \Lambda$ and all $t = 1, \dots, T$.

Pointwise maximization (for each (α^t, λ) tuple) of the integrand in ($\mathcal{Q}'$) while ignoring (for now) the remaining constraints yields the following solution:

$$q^*_t(\alpha^t, \lambda) = \max\left\{\varphi(\alpha^t, \lambda), 0\right\}.$$

It is important to note that (unlike the buyer's virtual value when values follow a recombinant binomial tree), the virtual value $\varphi(\alpha^t, \lambda)$ need not be path independent: without additional restrictions on the conditional distribution $G(\cdot|\lambda)$, there may be $\alpha^t, \hat{\alpha}^t \in A^t$ such that $v(\alpha^t) = v(\hat{\alpha}^t)$ but the summation in Equation (22) yields $\varphi(\alpha^t, \lambda) \neq \varphi(\hat{\alpha}^t, \lambda)$.[19] Meanwhile, our approach to solving for the optimal long-term contract (considering the single-deviation relaxation of the seller's problem) relies on pairing a path-independent allocation rule with a path-independent pricing rule to guarantee incentive compatibility with respect to compound deviations.

To enable this approach, we impose an additional separability assumption on the conditional distribution of shocks in order to guarantee path-independence of the allocation rule:

Assumption B.1. *There exist some constants $a, b \in \mathbb{R}$ and a function $\gamma : \Lambda \to \mathbb{R}$ such that*

$$\frac{\partial G(\alpha|\lambda)/\partial\lambda}{g(\alpha|\lambda)} = \alpha(a + b\log(\alpha))\gamma(\lambda)$$

for all $\alpha \in A$ and all $\lambda \in \Lambda$.

---

[18] Since values are multiplicative in the shocks, the additional distortions generated by the dependence of $v(\alpha^t)$ on λ have the additive structure above. This is in contrast to, for example, the nature of distortions in Besanko (1985), the autoregressive values examples of Pavan, Segal, and Toikka (2009, 2010), and the model of Krähmer and Strausz (2011).

[19] It is still true, however, that shocks commute: $\varphi(\alpha^t, \lambda) = \varphi(\sigma(\alpha^t), \lambda)$ for all $\alpha^t \in A^t$ and all possible permutations σ.

Notice that, under this assumption, we may write the buyer's virtual value as

$$\varphi(\alpha^t,\lambda) = v(\alpha^t) + \sum_{s=1}^{t} v(\alpha^t_{-s},\alpha^{s-1})\left[\alpha_s(a + b\log(\alpha_s)\gamma(\lambda)\right]\frac{1-F(\lambda)}{f(\lambda)}$$
$$= v(\alpha^t)\left[1 + \left(at + b\log(v(\alpha^t))\right)\gamma(\lambda)\frac{1-F(\lambda)}{f(\lambda)}\right],$$

where the second equality follows from the fact that $v(\alpha^t_{-s},\alpha^{s-1})\alpha_s = v(\alpha^t)$. Thus, Assumption B.1 implies that the period-t virtual value (and hence the allocation rule above) depends only on t, on λ, and on the buyer's *value* in that period, but not on the specific sequence of shocks generating that value. This does *not* imply, however, that the allocation rule is independent of t, as for any distinct t and t', $v(\alpha^t) = v(\alpha^{t'})$ implies that $\varphi(\alpha^t,\lambda) \neq \varphi(\alpha^{t'},\lambda)$ (as long as $a \neq 0$).

Clearly, Assumption B.1 involves some loss of generality. However, there are many natural and commonly used parametric classes of distributions for which the condition is satisfied. For example, if the G is a power distribution, so $G(\alpha|\lambda) = \alpha^\lambda$, then $\frac{\partial G(\alpha|\lambda)/\partial\lambda}{g(\alpha|\lambda)} = \frac{\alpha\log(\alpha)}{\lambda}$. Similarly, if G is an exponential distribution with mean λ, a Pareto distribution with minimum value λ and arbitrary shape parameter, or a truncated normal distribution with mean 0 and variance λ^2, then $\frac{\partial G(\alpha|\lambda)/\partial\lambda}{g(\alpha|\lambda)} = -\frac{\alpha}{\lambda}$. When the shocks are distributed according to a lognormal distribution with mean λ (and arbitrary non-zero variance), then the ratio in question equals $-\alpha$. Thus, while Assumption B.1 is *not* without loss of generality, it certainly does not rule out all cases of interest.

Therefore, under Assumption B.1, it is possible to write the optimal allocation rule as a function $\hat{q}^*_t$ of the buyer's reported value $v(\alpha^t)$ instead of the specific sequence of shocks α^t:

$$\hat{q}^*_t(v(\alpha^t),\lambda) := q^*_t(\alpha^t,\lambda) = \max\{\varphi(\alpha^t,\lambda),0\}. \tag{23}$$

We then pair this path-independent allocation rule with a path-independent payment rule that simply screens across each period's values as in a standard Mussa and Rosen (1978) nonlinear pricing problem. To this end, we define

$$p^*_t(\alpha^t,\lambda) := \hat{q}^*_t(v(\alpha^t),\lambda)v(\alpha^t) - \int_{\underline{\alpha}^t}^{v(\alpha^t)} \hat{q}^*_t(v',\lambda)\,\mathrm{d}v', \tag{24}$$

where $\underline{\alpha}^t$ is the buyer's lowest possible value in period-t. Thus, in each period t, the seller offers what is essentially a static screening mechanism $(q^*_t(\cdot,\lambda), p^*_t(\cdot,\lambda))$ that depends only on the initial report of λ. Note that Assumption B.1 implies that incentives for truthful reporting in the period-t mechanism are completely uncoupled from the incentives in any other period—the initial period report of λ determines specific menu offered in period t, but does not affect the buyer's incentives *within* that menu. Therefore, the single-deviation constraints (IC-t) are sufficient for "full" incentive compatibility. Standard results then yield the following necessary and sufficient condition for the proposed allocation rule in Equation (23) to satisfy these single-deviation constraints:

ASSUMPTION B.2. *For all $t = 1,\dots,T$, the allocation rule q^*_t that solves the seller's relaxed problem ($\mathcal{Q}'$) is (weakly) increasing in $v(\alpha^t)$.*

Moreover, note that—since $\varphi(\alpha^t, \lambda) \leq v(\alpha^t)$ for all $(\alpha^t, \lambda) \in A^t \times \Lambda$—the buyer's flow utility in each period (when reporting truthfully) is non-negative. This immediately implies that the period-t participation constraints (IR-t) are satisfied for all $t \geq 1$.

The final remaining piece of the optimal contract is the period-zero payment. However, since we have p_t^* for all $t \geq 1$, this payment is easily determined using the integral representation of $U_0(\lambda)$ from Lemma 5.2. In particular, note that Equation (16) implies that

$$p_0(\lambda) = \sum_{t=1}^{T} \delta^t \int_{A^t} (q_t(\alpha^t, \lambda) v(\alpha^t) - p_t(\alpha^t, \lambda)) \, \mathrm{d}W^t(\alpha^t|\lambda) - U_0(\lambda).$$

Therefore,

$$\begin{aligned} p_0^*(\lambda) := \sum_{t=1}^{T} \delta^t \int_{A^t} \int_{\underline{\alpha}^t}^{v(\alpha^t)} \hat{q}_t^*(v', \lambda) \, \mathrm{d}v' \, \mathrm{d}W^t(\alpha^t|\lambda) \\ + \sum_{t=1}^{T} \delta^t \int_{\underline{\lambda}}^{\lambda} \int_{A^t} \bar{q}_t^*(\alpha^t, \mu) \frac{\partial G(\alpha_t|\mu)/\partial\lambda}{g(\alpha_t|\mu)} \, \mathrm{d}W^t(\alpha^t|\mu) \, \mathrm{d}\mu. \end{aligned} \tag{25}$$

It remains to be seen that this contract is, in fact, incentive compatible, as the envelope condition derived in Lemma 5.2 is, in general, only a necessary condition for the initial-period single-deviation constraint (IC-0). As in Section 4, the additional assumption that the quantity schedules are increasing in λ *does* yield initial-period incentive compatibility.

ASSUMPTION B.3. *For all $t = 1, \ldots, T$, the allocation rule that solves the seller's relaxed problem ($\mathcal{Q}'$)—q_t^* in Equation (23)—is (weakly) increasing in λ.*

This assumption is the counterpart to Assumption A, and the following theorem (which we prove in the appendix) is the counterpart in this more general setting to Theorems 1 and 2.

THEOREM 3. *Suppose that Assumptions B.1, B.2, and B.3 are satisfied. Then the contract $(\mathbf{q}^*, \mathbf{p}^*)$, where $\mathbf{q}^*$ denotes the quantity schedules from Equation (23) and $\mathbf{p}^*$ denotes the payment rules from Equations (24) and (25), is an optimal contract that solves the seller's problem ($\mathcal{Q}'$).*

Having established this result, let us explore an important property of the optimal contract. Define $k_1(\lambda)$ to be the lowest value of α_1 that the buyer can report that, given her initial-period report of λ, leads to a non-negative allocation in period one; formally,

$$k_1(\lambda) := \inf \{\alpha_1' \in A : q_1^*(\alpha_1', \lambda) > 0\}. \tag{26}$$

The cutoff $k_1(\lambda)$ is decreasing (weakly) since, by Assumption B.3, q_1^* is increasing (weakly) in λ.

Similarly, define $k_t(\alpha^{t-1}, \lambda)$ to be the lowest value of α_t that the buyer can report in period t that, given her previous reports (α^{t-1}, λ), leads to a non-negative allocation in period t; that is, let

$$k_t(\alpha^{t-1}, \lambda) := \inf \left\{\alpha_t' \in A : q_t^*(\alpha_t', \alpha^{t-1}, \lambda) > 0\right\} \tag{27}$$

for all $t = 2, \ldots, T$, where we let $k_t(\alpha^{t-1}, \lambda) := \bar{\alpha}$ if the set above is empty. Again, because Assumptions B.2 and B.3 imply that q_t^* is (weakly) increasing in λ and α_s for all $s \leq t$, k_t is (weakly) decreasing in each of its arguments.

As $(\partial G(\bar{\alpha},\lambda)/\partial\lambda)/g(\bar{\alpha}|\lambda) = 0$ for all $\lambda \in \Lambda$, it must be the case that $q_1^*(\bar{\alpha},\lambda) = \bar{\alpha} > 0$ and thus $k_1(\lambda) < \bar{\alpha}$ for all λ. Similarly, this implies that $q_t(\bar{\alpha},\ldots,\bar{\alpha},\lambda) = \bar{\alpha}^t > 0$, and so (by continuity) $k_t(\bar{\alpha},\ldots,\bar{\alpha},\lambda) < \bar{\alpha}$ for all λ. Hence, for every possible initial-period report of λ, the set of reports leading to positive quantities in any period $t \geq 1$ is of non-zero measure. However, unlike $k_1(\lambda)$, it is entirely possible that $k_t(\alpha^{t-1},\lambda) = \bar{\alpha}$, implying that the buyer is not allocated any amount of the good, regardless of her period-t report.

One such instance is when $\alpha_t < k_t(\alpha^{t-1},\lambda)$ (or equivalently, whenever $q_t^*(\alpha^t,\lambda) = 0$). If this is the case, we must have

$$\varphi(\alpha^t,\lambda) = v(\alpha^t) + \sum_{s=1}^{t} v(\alpha^t_{-s},\alpha^{s-1}) \frac{\partial G(\alpha_s,\lambda)/\partial\lambda}{g(\alpha_s|\lambda)} \frac{1-F(\lambda)}{f(\lambda)} \leq 0.$$

But note that the buyer's period-$(t+1)$ virtual value, given α^t and λ, is

$$\varphi(\alpha_{t+1},\alpha^t,\lambda) = v(\alpha^{t+1}) + \sum_{s=1}^{t+1} v(\alpha^{t+1}_{-s},\alpha^{s-1}) \frac{\partial G(\alpha_s,\lambda)/\partial\lambda}{g(\alpha_s|\lambda)} \frac{1-F(\lambda)}{f(\lambda)}.$$

We may rewrite this as

$$\begin{aligned}\varphi(\alpha_{t+1},\alpha^t,\lambda) &= \alpha_{t+1}\left(v(\alpha^t) + \sum_{s=1}^{t} v(\alpha^t_{-s},\alpha^{s-1}) \frac{\partial G(\alpha_s,\lambda)/\partial\lambda}{g(\alpha_s|\lambda)} \frac{1-F(\lambda)}{f(\lambda)} \right) \\ &\quad + v(\alpha^t)\frac{\partial G(\alpha_{t+1}|\lambda)/\partial\lambda}{g(\alpha_{t+1}|\lambda)} \frac{1-F(\lambda)}{f(\lambda)} \\ &= \alpha_{t+1}\varphi(\alpha^t,\lambda) + v(\alpha^t)\frac{\partial G(\alpha_{t+1}|\lambda)/\partial\lambda}{g(\alpha_{t+1}|\lambda)} \frac{1-F(\lambda)}{f(\lambda)} \leq 0,\end{aligned}$$

where the inequality is guaranteed by the fact that $\partial G(\alpha|\lambda)/\partial\lambda \leq 0$ for all α and λ via first-order stochastic dominance. Therefore, $k_{t+1}(\alpha^t,\lambda) = \bar{\alpha}$ whenever $\alpha_t \leq k_t(\alpha^{t-1},\lambda)$. Thus, if the buyer is excluded in some period t, she continues to be excluded in *all* future periods, regardless of her reported shocks—once a buyer has been "cut off," she is cut off permanently.

In addition, note that the optimal contract involves a form of "tightening the screws," as the set of reports that lead to a positive quantity in any period $t+1 \leq T$ is contained in the corresponding set for period t. To see this, let $\alpha^* := k_t(\alpha^{t-1},\lambda)$, and suppose that $\alpha^* > \underline{\alpha}$, so that there are some period-t reports that lead to the buyer's exclusion. If the buyer's period-t shock is α^* (so she is "just barely" excluded), while her period-$(t+1)$ shock is any $\alpha > \alpha^*$, then

$$\begin{aligned}\varphi(\alpha,\alpha^*,\alpha^{t-1},\lambda) &= \alpha\varphi(\alpha^*,\alpha^{t-1},\lambda) + \alpha^* v(\alpha^{t-1})\frac{\partial G(\alpha|\lambda)/\partial\lambda}{g(\alpha|\lambda)} \frac{1-F(\lambda)}{f(\lambda)} \\ &= \alpha^* v(\alpha^{t-1})\frac{\partial G(\alpha|\lambda)/\partial\lambda}{g(\alpha|\lambda)} \frac{1-F(\lambda)}{f(\lambda)} \leq 0,\end{aligned}$$

where the second equality follows from the definition of α^* as the largest excluded shock in period-t (so $\varphi(\alpha^*,\alpha^{t-1},\lambda) = 0$). Therefore,

$$q^*_{t+1}(\alpha,\alpha^*,\alpha^{t-1},\lambda) = \max\left\{\alpha^* v(\alpha^{t-1})\frac{\partial G(\alpha|\lambda)/\partial\lambda}{g(\alpha|\lambda)} \frac{1-F(\lambda)}{f(\lambda)}, 0\right\} = 0,$$

implying that

$$k_{t+1}(k_t(\alpha^{t-1},\lambda),\alpha^{t-1},\lambda)=\bar{\alpha}.$$

Thus, a buyer who is on the cusp of allocation in period t is excluded in period $t+1$, regardless of her realized shock.

Meanwhile, note that for any $\alpha^t \in A^t$ and $\lambda \in \Lambda$, the virtual value of a buyer who experiences the highest possible shock $\bar{\alpha}$ is simply the product of this shock with her virtual value from the previous period; that is,

$$\varphi(\bar{\alpha},\alpha^t,\lambda)=\bar{\alpha}\varphi(\alpha^t,\lambda)+v(\alpha^t)\frac{\partial G(\bar{\alpha}|\lambda)/\partial\lambda}{g(\bar{\alpha}|\lambda)}\frac{1-F(\lambda)}{f(\lambda)}=\bar{\alpha}\varphi(\alpha^t,\lambda).$$

This expression is positive if, and only if, $\alpha_t \geq k_t(\alpha^{t-1},\lambda)$; thus,

$$k_{t+1}(\bar{\alpha},\alpha^{t-1},\lambda)=k_t(\alpha^{t-1},\lambda).$$

Thus, if a buyer receives a positive quantity of the good in period t and then observes the highest possible shock $\bar{\alpha}$ in period $t+1$, she will again be allocated a positive quantity period $t+1$. In fact, a straightforward continuity argument implies that a buyer allocated a positive quantity in period t will be allocated a positive quantity in period $t+1$ as long as her shock is sufficiently close to $\bar{\alpha}$.

Finally, recall that $k_{t+1}(\alpha^t,\lambda)$ is decreasing in α_t. This property, combined with the two observations above, implies that, for any $\alpha_t \geq k_t(\alpha^{t-1},\lambda)$, the set of "admissible" period-$(t+1)$ reports $[k_{t+1}(\alpha^t,\lambda),\bar{\alpha}]$ that lead to a positive allocation in period $t+1$ is a subset of the corresponding set of "admissible" period-t reports $[k_t(\alpha^{t-1},\lambda),\bar{\alpha}]$.

This feature of the optimal contract is the continuous analog of the finite honeymoon phases in the discrete shock single-unit demand setting of Section 4.1. Recall that the optimal contract in that setting allowed, for each initial-period report λ, a fixed number of "low" d reports before excluding the buyer from future allocations, implying that the probability of contract termination by the seller was increasing over time. In the continuous-convex setting considered here, this effect is captured by the fact that the set of reports that lead to an allocation (or, equivalently, that prevent permanent exclusion) is shrinking over time. Thus, as in Section 4, the seller progressively screens the buyer by restricting supply and increasing the probability of permanent exclusion as the relationship progresses.

6. Conclusion

In this paper, we examine a model of long-term contracting in which the buyer is not only privately informed about her value at every point in time, but also about the process by which her value evolves. We then solve for the seller's optimal contract, taking into account the buyer's incentives for participation and for truthful revelation throughout the interaction. As in earlier work in dynamic mechanism design and sequential screening, we demonstrate that option contracts—contracts where the seller offers the buyer her choice from a menu of "buy" options—are profit-maximizing mechanisms. However, by extending the literature to more general time horizons, we are able to study the dynamic features of these contracts. Thus, the optimal long-term contract in our more general setting features surprisingly simple menus of options that vary not only by up-front cost and future strike price, but also by the generosity of quantity provision over the course of the contract. In particular, these more generous choices require greater upfront investments by the buyer while featuring lower strike prices. Moreover, we identify an additional mechanism by which the seller price discriminates across buyers with differing willingness to pay: over time, sales are made to fewer and fewer buyers, as the seller progressively screens and excludes lower-valued buyers and ratchets prices upwards, thereby reducing the rents paid to higher-valued buyers. In the long-run, this leads to inefficiently early termination of the buyer-seller relationship, a feature that is not readily intuited using the two-period models prevalent in the literature.

Our model and results set the stage for several avenues of further inquiry. For example, there are a number of settings where the contracting environment or the value of the relationship are influenced by investments made by the agent. Exploring the dynamics of contracting in such an environment would advance our understanding of incentive provision beyond the present work's focus on adverse selection. Another set of interesting questions arise when considering the seller's power to commit: since the optimal contract in our setting is *not* renegotiation-proof, our assumption of full commitment power has substantial bite. Understanding the precise role of commitment would therefore be a natural topic for additional investigation. Finally, competition among both buyers and sellers in a dynamic environment such as our own is not particularly well-understood; progress in this direction would greatly advance our knowledge and yield important insights for market analysis and design. We leave these questions, however, for future research.

APPENDIX OF OMITTED PROOFS

PROOF OF LEMMA 3.1. Notice that the single-deviation constraint (IC-t) may be rewritten as

$$U_t(\alpha_t, \alpha^{t-1}, \lambda) \geq q_t(\alpha'_t, \alpha^{t-1}, \lambda)v(\alpha^t) - p_t(\alpha'_t, \alpha^{t-1}, \lambda)$$
$$+ \sum_{s=t+1}^{T} \delta^{s-t}\mathbb{E}\left[q_s(\alpha^s_{-t}, \alpha'_t, \alpha^{t-1}, \lambda)v(\alpha^s_{-t}, \alpha^t) - p_s(\alpha^s_{-t}, \alpha'_t, \alpha^{t-1}, \lambda)\middle|\alpha^t, \lambda\right].$$

Adding and subtracting the quantity

$$q_t(\alpha'_t, \alpha^{t-1}, \lambda)v(\alpha'_t, \alpha^{t-1}) + \sum_{s=t+1}^{T} \delta^{s-t}\mathbb{E}\left[q_s(\alpha^s_{-t}, \alpha'_t, \alpha^{t-1}, \lambda)v(\alpha^s_{-t}, \alpha'_t, \alpha^{t-1})\middle|\alpha^t, \lambda\right]$$

from the right-hand side of the inequality yields

$$U_t(\alpha_t, \alpha^{t-1}, \lambda) \geq U_t(\alpha'_t, \alpha^{t-1}, \lambda) + (\alpha_t - \alpha'_t)q_t(\alpha'_t, \alpha^{t-1}, \lambda)v(\alpha^{t-1})$$
$$+ (\alpha_t - \alpha'_t) \sum_{s=t+1}^{\infty} \delta^{s-t}\mathbb{E}\left[q_s(\alpha^s_{-t}, \alpha'_t, \alpha^{t-1}, \lambda)v(\alpha^s_{-t}, \alpha^{t-1})\middle|\alpha^t, \lambda\right].$$

Therefore, for all $\alpha_t, \alpha'_t \in A$, $\alpha^{t-1} \in A^{t-1}$ and $\lambda \in \Lambda$, we have

$$U_t(\alpha_t, \alpha^{t-1}, \lambda) \geq U_t(\alpha'_t, \alpha^{t-1}, \lambda) + (\alpha_t - \alpha'_t)\bar{q}_t(\alpha'_t, \alpha^{t-1}, \lambda).$$

Letting $\alpha_t = u$ and $\alpha'_t = d$, we can write the inequality above as

$$U_t(u, \alpha^{t-1}, \lambda) \geq U_t(d, \alpha^{t-1}, \lambda) + \bar{q}_t(d, \alpha^{t-1}, \lambda)\Delta \text{ for all } \alpha^{t-1} \in A^{t-1} \text{ and } \lambda \in \Lambda,$$

where $\Delta := u - d$. In addition, letting $\alpha_t = d$ and $\alpha'_t = u$, the inequality above implies that

$$U_t(d, \alpha^{t-1}, \lambda) \geq U_t(u, \alpha^{t-1}, \lambda) - \bar{q}_t(u, \alpha^{t-1}, \lambda)\Delta \text{ for all } \alpha^{t-1} \in A^{t-1} \text{ and } \lambda \in \Lambda.$$

Notice that rearranging the first of these two inequalities immediately yields condition (IC′-t). Similarly, adding the two inequalities yields condition (MON-t).

Finally, with conditions (IC′-t) and (MON-t) in hand, (IR-t) is satisfied only if

$$U_t(d, \alpha^{t-1}, \lambda) \geq 0 \text{ for all } \alpha^{t-1} \in A^{t-1} \text{ and } \lambda \in \Lambda;$$

that is, only if (IR′-t) holds.

Note that the sufficiency of the conditions derived above for the period-t single-deviation and individual rationality constraints follows immediately via basic arithmetic. □

PROOF OF LEMMA 3.2. Note first that, for all $s \leq T$, the definition of $U_s(\alpha^s, \lambda')$ implies that

$$\sum_{t=s}^{T} \delta^t \mathbb{E}\left[q_t(\alpha^t, \lambda')v(\alpha^t) - p_t(\alpha^t, \lambda')\middle|\lambda\right]$$
$$= \delta^s \mathbb{E}\left[U_s(\alpha^s, \lambda')\middle|\lambda\right] + \sum_{t=s+1}^{T} \delta^t \mathbb{E}\left[q_t(\alpha^t, \lambda')v(\alpha^t) - p_t(\alpha^t, \lambda')\middle|\lambda\right]$$
$$- \sum_{t=s+1}^{T} \delta^t \mathbb{E}\left[\mathbb{E}\left[q_t(\alpha^t_{-s}, \alpha^s, \lambda')v(\alpha^t_{-s}, \alpha^s) - p_t(\alpha^t_{-s}, \alpha^s, \lambda')\middle|\alpha^s, \lambda'\right]\middle|\lambda\right].$$

Also, note that for all $s \leq T$,

$$
\begin{aligned}
&\sum_{t=s+1}^{T} \delta^t \mathbb{E}\left[\mathbb{E}\left[q_t(\alpha_{-s}^t, \alpha^s, \lambda')v(\alpha_{-s}^t, \alpha^s) - p_t(\alpha_{-s}^t, \alpha^s, \lambda')\middle|\alpha^s, \lambda'\right]\middle|\lambda\right] \\
&\qquad = \mathbb{E}\left[\sum_{t=s+1}^{T} \delta^t \mathbb{E}\left[q_t(\alpha_{-s}^t, \alpha^s, \lambda')v(\alpha_{-s}^t, \alpha^s) - p_t(\alpha_{-s}^t, \alpha^s, \lambda')\middle|\alpha^s, \lambda'\right]\middle|\lambda\right] \\
&\qquad = \delta^{s+1}\mathbb{E}\left[\mathbb{E}\left[U_{s+1}(\alpha_{s+1}, \alpha^s, \lambda')\middle|\alpha^s, \lambda'\right]\middle|\lambda\right],
\end{aligned}
$$

where we have made use of the definition of $U_{s+1}(\alpha^s, \lambda')$. With this in hand, recall the definition of $\hat{U}_0(\lambda', \lambda)$ from Equation (2):

$$
\hat{U}_0(\lambda', \lambda) := -p_0(\lambda') + \sum_{t=1}^{T} \delta^t \mathbb{E}\left[q_t(\alpha^t, \lambda')v(\alpha^t) - p_t(\alpha^t, \lambda')\middle|\lambda\right].
$$

We can rewrite this expression as

$$
\begin{aligned}
\hat{U}_0(\lambda', \lambda) &= -p_0(\lambda') + \delta\mathbb{E}\left[U_1(\alpha_1, \lambda')\middle|\lambda\right] + \sum_{t=2}^{T} \delta^t \mathbb{E}\left[q_t(\alpha^t, \lambda')v(\alpha^t) - p_t(\alpha^t, \lambda')\middle|\lambda\right] \\
&\quad - \sum_{t=2}^{T} \delta^t \mathbb{E}\left[\mathbb{E}\left[q_t(\alpha_{-1}^t, \alpha_1, \lambda')v(\alpha_{-1}^t, \alpha_1) - p_t(\alpha_{-1}^t, \alpha_1, \lambda')\middle|\alpha_1, \lambda'\right]\middle|\lambda\right] \\
&= -p_0(\lambda') + \delta\mathbb{E}\left[U_1(\alpha_1, \lambda')\middle|\lambda\right] + \delta^2\mathbb{E}\left[U_2(\alpha^2, \lambda')\middle|\lambda\right] - \delta^2\mathbb{E}\left[\mathbb{E}\left[U_2(\alpha_2, \alpha_1, \lambda')\middle|\alpha_1, \lambda'\right]\middle|\lambda\right] \\
&\quad + \sum_{t=3}^{T} \delta^t \mathbb{E}\left[q_t(\alpha^t, \lambda')v(\alpha^t) - p_t(\alpha^t, \lambda')\middle|\lambda\right] \\
&\quad - \sum_{t=3}^{T} \delta^t \mathbb{E}\left[\mathbb{E}\left[q_t(\alpha_{-2}^t, \alpha^2, \lambda')v(\alpha_{-2}^t, \alpha^2) - p_t(\alpha_{-2}^t, \alpha^2, \lambda')\middle|\alpha^2, \lambda'\right]\middle|\lambda\right]
\end{aligned}
$$

Proceeding inductively, we may conclude that

$$
\begin{aligned}
\hat{U}_0(\lambda', \lambda) &= -p_0(\lambda') + \delta\mathbb{E}\left[U_1(\alpha_1, \lambda')\middle|\lambda\right] \\
&\quad + \sum_{t=2}^{T} \delta^t \mathbb{E}\left[\mathbb{E}\left[U_t(\alpha_t, \alpha^{t-1}, \lambda')\middle|\alpha^{t-1}, \lambda\right] - \mathbb{E}\left[U_t(\alpha_t, \alpha^{t-1}, \lambda')\middle|\alpha^{t-1}, \lambda'\right]\middle|\lambda\right] \\
&= -p_0(\lambda') + \delta\left(\lambda\left(U_1(u, \lambda') - U_1(d, \lambda')\right) + U_1(d, \lambda')\right) \\
&\quad + (\lambda - \lambda')\sum_{t=2}^{T} \delta^t \mathbb{E}\left[U_t(u, \alpha^{t-1}, \lambda') - U_t(d, \alpha^{t-1}, \lambda')\middle|\lambda\right].
\end{aligned}
$$

With this in hand, note that

$$
\begin{aligned}
\frac{\partial}{\partial \lambda}\hat{U}_0(\lambda', \lambda) &= \sum_{t=1}^{T} \delta^t \mathbb{E}\left[U_t(u, \alpha^{t-1}, \lambda') - U_t(d, \alpha^{t-1}, \lambda')\middle|\lambda\right] \\
&\quad + (\lambda - \lambda')\sum_{t=2}^{T} \delta^t \frac{\partial}{\partial \lambda}\left(\mathbb{E}\left[U_t(u, \alpha^{t-1}, \lambda') - U_t(d, \alpha^{t-1}, \lambda')\middle|\lambda\right]\right).
\end{aligned}
$$

Since condition (IC-0) requires that $\hat{U}_0(\lambda,\lambda) = \max_{\lambda'}\{\hat{U}_0(\lambda',\lambda)\}$ for all λ, the envelope theorem (see Milgrom and Segal (2002)) implies that

$$U_0'(\lambda) = \frac{\partial}{\partial\lambda}\hat{U}_0(\lambda',\lambda)\bigg|_{\lambda'=\lambda} = \sum_{t=1}^{T}\delta^t\mathbb{E}\left[U_t(u,\alpha^{t-1},\lambda) - U_t(d,\alpha^{t-1},\lambda)\Big|\lambda\right]. \qquad \square$$

PROOF OF THEOREM 1. Note that we may rewrite $\hat{U}_0(\lambda',\lambda)$ from Equation (2) as

$$\hat{U}_0(\lambda',\lambda) = \int_0^{\lambda'}\sum_{t=1}^{T}\delta^t\mathbb{E}\left[\bar{q}_t^*(d,\alpha^{t-1},\mu)\Delta\Big|\mu\right]\,\mathrm{d}\mu - \sum_{t=1}^{T}\delta^t\mathbb{E}\left[q_t^*(\alpha^t,\lambda')(v(\alpha^t)-p_t^*(\alpha^t,\lambda'))\big|\lambda'\right]$$
$$+\sum_{t=1}^{T}\delta^t\mathbb{E}\left[q_t^*(\alpha^t,\lambda')(v(\alpha^t)-p_t^*(\alpha^t,\lambda'))\big|\lambda\right].$$

Therefore,

$$\hat{U}_0(\lambda,\lambda) - \hat{U}_0(\lambda',\lambda) = \int_{\lambda'}^{\lambda}\sum_{t=1}^{T}\delta^t\mathbb{E}\left[\bar{q}_t^*(d,\alpha^{t-1},\mu)\Delta\Big|\mu\right]\,\mathrm{d}\mu - \sum_{t=1}^{T}\delta^t\mathbb{E}\left[q_t^*(\alpha^t,\lambda')(v(\alpha^t)-p_t^*(\alpha^t,\lambda'))\big|\lambda\right]$$
$$+\sum_{t=1}^{T}\delta^t\mathbb{E}\left[q_t^*(\alpha^t,\lambda')(v(\alpha^t)-p_t^*(\alpha^t,\lambda'))\big|\lambda'\right].$$

Since $q_t^*(\alpha^t,\cdot)$ is non-decreasing for all t and α^t (due to Assumption A), so is $\bar{q}_t^*(d,\alpha^{t-1},\cdot)$. Therefore,

$$\begin{aligned}\hat{U}_0(\lambda,\lambda) - \hat{U}_0(\lambda',\lambda) &\geq \int_{\lambda'}^{\lambda}\sum_{t=1}^{T}\delta^t\mathbb{E}\left[\bar{q}_t^*(d,\alpha^{t-1},\lambda')\Delta\Big|\mu\right]\,\mathrm{d}\mu \\ &\quad+\sum_{t=1}^{T}\delta^t\mathbb{E}\left[q_t^*(\alpha^t,\lambda')(v(\alpha^t)-p_t^*(\alpha^t,\lambda'))\big|\lambda'\right] \\ &\quad-\sum_{t=1}^{T}\delta^t\mathbb{E}\left[q_t^*(\alpha^t,\lambda')(v(\alpha^t)-p_t^*(\alpha^t,\lambda'))\big|\lambda\right] \\ &= \int_{\lambda'}^{\lambda}\sum_{t=1}^{T}\sum_{s=1}^{t}\delta^t\mathbb{E}\left[q_t^*(\alpha_{-s}^t,d,\alpha^{s-1},\lambda')v(\alpha_{-s}^t,\alpha^{s-1})\Delta\Big|\mu\right]\,\mathrm{d}\mu \\ &\quad+\sum_{t=1}^{T}\delta^t\mathbb{E}\left[q_t^*(\alpha^t,\lambda')(v(\alpha^t)-p_t^*(\alpha^t,\lambda'))\big|\lambda'\right] \\ &\quad-\sum_{t=1}^{T}\delta^t\mathbb{E}\left[q_t^*(\alpha^t,\lambda')(v(\alpha^t)-p_t^*(\alpha^t,\lambda'))\big|\lambda\right],\end{aligned}$$

where the equality follows from the identity in Equation (6).

For each $t = 1,\ldots,T$, let $m_t := k_t(\lambda')$, and note that for all $\mu\in\Lambda$, we have

$$\begin{aligned}\mathbb{E}\left[q_t^*(\alpha^t,\lambda')(v(\alpha^t)-p_t^*(\alpha^t,\lambda'))\big|\mu\right] &= \sum_{j=0}^{\min\{m_t,t\}}\binom{t}{j}\mu^{t-j}(1-\mu)^j\left(u^{t-j}d^j - u^{t-\min\{m_t,t\}}d^{\min\{m_t,t\}}\right) \\ &= \begin{cases}\sum_{j=0}^{t}\binom{t}{j}\mu^{t-j}(1-\mu)^j\left(u^{t-j}d^j - d^t\right) & \text{if } m_t\geq t, \\ \sum_{j=0}^{m_t}\binom{t}{j}\mu^{t-j}(1-\mu)^j\left(u^{t-j}d^j - u^{t-m_t}d^{m_t}\right) & \text{if } m_t < t.\end{cases}\end{aligned}$$

Therefore, we may write (for each $t = 1, \dots, T$)

$$\begin{aligned}
&\mathbb{E}\left[q_t^*(\alpha^t, \lambda')(v(\alpha^t) - p_t^*(\alpha^t, \lambda'))\middle|\lambda'\right] - \mathbb{E}\left[q_t^*(\alpha^t, \lambda')(v(\alpha^t) - p_t^*(\alpha^t, \lambda'))\middle|\lambda\right] \\
&\qquad = \begin{cases} \sum_{j=0}^{t} \binom{t}{j} \left[\mu^{t-j}(1-\mu)^j\right]_{\mu=\lambda}^{\lambda'} u^{t-j} d^j & \text{if } m_t \geq t, \\ \sum_{j=0}^{m_t-1} \binom{t}{j} \left[\mu^{t-j}(1-\mu)^j\right]_{\mu=\lambda}^{\lambda'} u^{t-j} d^j & \\ \quad - \sum_{j=0}^{m_t-1} \binom{t}{j} \left[\mu^{t-j}(1-\mu)^j\right]_{\mu=\lambda}^{\lambda'} u^{t-m} d^m & \text{if } m_t < t. \end{cases}
\end{aligned}$$

Meanwhile, note that for each $t = 1, \dots, T$,

$$\begin{aligned}
\sum_{s=1}^{t} \mathbb{E}\left[q_t^*(\alpha_{-s}^t, d, \alpha^{s-1}, \lambda') v(\alpha_{-s}^t, \alpha^{s-1})\Delta \middle| \mu\right] &= \sum_{s=1}^{t} \sum_{j=0}^{\min\{m_t, t\}} \binom{t-1}{j} \mu^{t-1-j}(1-\mu)^j u^{t-1-j} d^j \Delta \\
&= \sum_{j=0}^{\min\{m_t, t\}} t \binom{t-1}{j} \mu^{t-1-j}(1-\mu)^j u^{t-1-j} d^j \Delta,
\end{aligned}$$

so we must have

$$\begin{aligned}
&\sum_{s=1}^{t} \mathbb{E}\left[q_t^*(\alpha_{-s}^t, d, \alpha^{s-1}, \lambda') v(\alpha_{-s}^t, \alpha^{s-1})\Delta \middle| \mu\right] \\
&\qquad = \begin{cases} t(\mu\Delta + d)^{t-1}\Delta & \text{if } m_t \geq t, \\ \sum_{j=0}^{m_t} t\binom{t-1}{j} \mu^{t-1-j}(1-\mu)^j u^{t-1-j} d^j \Delta & \text{if } m_t < t. \end{cases}
\end{aligned}$$

This implies that, for all t such that $m_t \geq t$,

$$\begin{aligned}
\int_{\lambda'}^{\lambda} \sum_{s=1}^{t} \mathbb{E}\left[q_t^*(\alpha_{-s}^t, d, \alpha^{s-1}, \lambda') v(\alpha_{-s}^t, \alpha^{s-1})\Delta \middle| \mu\right] \, \mathrm{d}\mu &= \int_{\lambda'}^{\lambda} t(\mu\Delta + d)^{t-1}\Delta \, \mathrm{d}\mu \\
&= \sum_{j=0}^{t} \binom{t}{j} \left[\mu^{t-j}(1-\mu)^j\right]_{\mu=\lambda'}^{\lambda} u^{t-j} d^j.
\end{aligned}$$

Meanwhile, for all t such that $m_t < t$, we may write

$$\begin{aligned}
&\sum_{j=0}^{m_t} t\binom{t-1}{j} \mu^{t-1-j}(1-\mu)^j u^{t-1-j} d^j \Delta \\
&\qquad = \sum_{j=0}^{m_t-1} t \left[\binom{t-1}{j} \mu^{t-j-1}(1-\mu)^j - \binom{t-1}{j-1} \mu^{t-j}(1-\mu)^{j-1}\right] u^{t-j} d^j \\
&\qquad\quad - t\binom{t-1}{m_t-1} \mu^{t-m_t}(1-\mu)^{m_t-1} u^{t-m_t} d^{m_t} \\
&\qquad = \sum_{j=0}^{m_t-1} \binom{t}{j} \left[(t-j)\mu^{t-j-1}(1-\mu)^j - j\mu^{t-j}(1-\mu)^{j-1}\right] u^{t-j} d^j \\
&\qquad\quad - t\binom{t-1}{m_t-1} u^{t-m_t} d^{m_t} \sum_{k=0}^{m_t-1} \binom{m_t-1}{k} (-1)^k \mu^{t-m_t+k}.
\end{aligned}$$

Thus, for all t with $m_t < t$, we have

$$\int_{\lambda'}^{\lambda} \sum_{s=1}^{t} \mathbb{E}\left[q_t^*(\alpha_{-s}^t, d, \alpha^{s-1}, \lambda') v(\alpha_{-s}^t, \alpha^{s-1}) \Delta \middle| \mu \right] \mathrm{d}\mu$$

$$= \binom{t}{j} u^{t-j} d^j \int_{\lambda'}^{\lambda} \left[(t-j)\mu^{t-j-1}(1-\mu)^j - j\mu^{t-j}(1-\mu)^{j-1} \right] \mathrm{d}\mu$$

$$- t\binom{t-1}{m_t-1} u^{t-m_t} d^{m_t} \int_{\lambda'}^{\lambda} \sum_{k=0}^{m_t-1} \binom{m_t-1}{k} (-1)^k \mu^{t-m_t+k}$$

$$= \binom{t}{j} \left[\mu^{t-j}(1-\mu)^j \right]_{\mu=\lambda'}^{\lambda} u^{t-j} d^j$$

$$- \sum_{k=0}^{m_t-1} \left[\frac{t}{t-m_t+k+1} \binom{t-1}{m_t-1} \binom{m_t-1}{k} (-1)^k \mu^{t-m_t+k+1} \right]_{\mu=\lambda'}^{\lambda} u^{t-m_t} d^{m_t}.$$

Finally, note that

$$\frac{t}{t-m_t+k+1} \binom{t-1}{m_t-1} \binom{m_t-1}{k} (-1)^k \mu^{t-m_t+k+1}$$

$$= \frac{t-m_t+1}{t-m_t+k+1} \binom{t}{m_t-1} \binom{m_t-1}{k} (-1)^k \mu^{t-m_t+k+1}$$

$$= \binom{t}{m_t-1-k} \binom{t-m_t+k}{k} (-1)^k \mu^{t-m_t+k+1}.$$

Using the binomial identity $\binom{n-1}{k}(-1)^k = \sum_{j=0}^{k} \binom{n}{j}(-1)^j$, we may write

$$\sum_{k=0}^{m_t-1} \frac{t}{t-m_t+k+1} \binom{t-1}{m_t-1} \binom{m_t-1}{k} (-1)^k \mu^{t-m_t+k+1}$$

$$= \sum_{k=0}^{m_t-1} \sum_{j=0}^{k} \binom{t}{j+m_t-k-1} \binom{j+m_t-k-1}{j} (-1)^j \mu^{t-m_t+k+1}$$

$$= \sum_{j=0}^{m_t-1} \sum_{k=0}^{j} \binom{t}{j} \binom{j}{k} (-1)^k \mu^{t-j+k}$$

$$= \sum_{j=0}^{m_t-1} \binom{t}{j} \mu^{t-j}(1-\mu)^j.$$

Therefore, for each $t = 1, \ldots, T$, we may conclude that

$$\int_{\lambda'}^{\lambda} \mathbb{E}\left[\bar{q}_t^*(d, \alpha^{t-1}, \lambda') \Delta \middle| \mu \right] \mathrm{d}\mu$$

$$+ \mathbb{E}\left[q_t^*(\alpha^t, \lambda')(v(\alpha^t) - p_t^*(\alpha^t, \lambda')) \middle| \lambda' \right] - \mathbb{E}\left[q_t^*(\alpha^t, \lambda')(v(\alpha^t) - p_t^*(\alpha^t, \lambda')) \middle| \lambda \right] = 0.$$

Therefore, for all $\lambda, \lambda' \in \Lambda$, $\hat{U}_0(\lambda, \lambda) \geq \hat{U}_0(\lambda', \lambda)$; that is, for each $\lambda \in \Lambda$, $\hat{U}_0(\lambda', \lambda)$ achieves a global maximum when $\lambda' = \lambda$, implying that the buyer has no incentive to misreport her private information in the initial-period. Combined with the observation that the mechanism $(\mathbf{q}^*, \mathbf{p}^*)$ is incentive compatible in all $t \geq 1$, this implies that this mechanism does, in fact, maximize the seller's profits. □

Proof of Theorem 2. Recall from Equation (2) that we may write $\hat{U}_0(\lambda', \lambda)$ as

$$\begin{aligned}\hat{U}_0(\lambda', \lambda) = &\int_0^{\lambda'} \sum_{t=1}^{T} \delta^t \mathbb{E}\left[\bar{q}_t^*(d, \alpha^{t-1}, \mu)\Delta \middle| \mu\right] \mathrm{d}\mu - \sum_{t=1}^{T} \delta^t \mathbb{E}\left[q_t^*(\alpha^t, \lambda')(v(\alpha^t) - p_t^*(\alpha^t, \lambda'))\middle|\lambda'\right] \\ &+ \sum_{t=1}^{T} \delta^t \mathbb{E}\left[q_t^*(\alpha^t, \lambda')(v(\alpha^t) - p_t^*(\alpha^t, \lambda'))\middle|\lambda\right] \\ = &\int_0^{\lambda'} \sum_{t=1}^{T} \delta^t \sum_{s=1}^{t} \mathbb{E}\left[q_s^*(\alpha_{-s}^t, d, \alpha^{s-1}, \mu) v(\alpha_{-s}^t, \alpha^{s-1})\Delta \middle| \mu\right] \mathrm{d}\mu \\ &- \sum_{t=1}^{T} \delta^t \mathbb{E}\left[q_t^*(\alpha^t, \lambda')(v(\alpha^t) - p_t^*(\alpha^t, \lambda'))\middle|\lambda'\right] \\ &+ \sum_{t=1}^{T} \delta^t \mathbb{E}\left[q_t^*(\alpha^t, \lambda')(v(\alpha^t) - p_t^*(\alpha^t, \lambda'))\middle|\lambda\right],\end{aligned}$$

where the equality comes from the identity in Equation (6). Since for all t and all $\mu \in \Lambda$, $q_t^*(\alpha^t, \mu)$ only depends on α^t through $m(\alpha^t) = \sum_{s=1}^{t} \mathbb{1}_d(\alpha_s)$, we will abuse notation slightly and write $q_t^*(k, \mu)$ to denote the quantity allocated in period t to a buyer who has reported (α^t, μ) with $m(\alpha^t) = k$. Therefore, we can rewrite the expression above as

$$\begin{aligned}\hat{U}_0(\lambda', \lambda) = &\int_0^{\lambda'} \sum_{t=1}^{T} \delta^t \sum_{k=0}^{t-1} t \binom{t-1}{k} \mu^{t-1-k}(1-\mu)^k q_t^*(k+1, \mu)\Delta u^{t-1-k} d^k \, \mathrm{d}\mu \\ &- \sum_{t=1}^{T} \delta^t \sum_{k=0}^{t} \binom{t}{k} (\lambda')^{t-k}(1-\lambda')^k \sum_{j=k+1}^{t} q_t^*(j, \lambda')\Delta u^{t-j} d^{j-1} \\ &+ \sum_{t=1}^{T} \delta^t \sum_{k=0}^{t} \binom{t}{k} \lambda^{t-k}(1-\lambda)^k \sum_{j=k+1}^{t} q_t^*(j, \lambda')\Delta u^{t-j} d^{j-1}.\end{aligned}$$

Taking the partial derivative of the expression above with respect to λ' yields

$$\begin{aligned}&\frac{\partial \hat{U}_0(\lambda', \lambda)}{\partial \lambda'} \\ &= \sum_{t=1}^{T} \delta^t \sum_{k=0}^{t-1} t \binom{t-1}{k} (\lambda')^{t-1-k}(1-\lambda')^k q_t^*(k+1, \lambda')\Delta u^{t-1-k} d^k \\ &\quad - \sum_{t=1}^{T} \delta^t \sum_{k=0}^{t} \binom{t}{k} \left((t-k)(\lambda')^{t-k-1}(1-\lambda')^k - k(\lambda')^{t-k}(1-\lambda')^{k-1}\right) \sum_{j=k+1}^{t} q_t^*(j, \lambda')\Delta u^{t-j} d^{j-1} \\ &\quad + \sum_{t=1}^{T} \delta^t \sum_{k=0}^{t} \binom{t}{k} \left(\lambda^{t-k}(1-\lambda)^k - (\lambda')^{t-k}(1-\lambda')^k\right) \sum_{j=k+1}^{t} \frac{\partial q_t^*(j, \lambda')}{\partial \lambda'} \Delta u^{t-j} d^{j-1}.\end{aligned}$$

Fix an arbitrary $t \geq 1$, and note that (ignoring the δ^t coefficient) the summand in the second line of the expression above may be rewritten as

$$\sum_{k=0}^{t-1} \binom{t}{k} \left((t-k)(\lambda')^{t-k-1}(1-\lambda')^k - k(\lambda')^{t-k}(1-\lambda')^{k-1}\right) \sum_{j=k+1}^{t} q_t^*(j, \lambda')\Delta u^{t-j} d^{j-1},$$

where we have used the fact that the innermost (rightmost) summation equals zero when $k = t$. Reversing the order of summation, this quantity becomes

$$\sum_{j=1}^{t}\sum_{k=0}^{j-1}\binom{t}{k}\left((t-k)(\lambda')^{t-k-1}(1-\lambda')^{k}-k(\lambda')^{t-k}(1-\lambda')^{k-1}\right)q_t^*(j,\lambda')\Delta u^{t-j}d^{j-1}$$
$$=\sum_{j=0}^{t-1}q_t^*(j+1,\lambda')\Delta u^{t-j-1}d^{j}\sum_{k=0}^{j}\binom{t}{k}\left((t-k)(\lambda')^{t-k-1}(1-\lambda')^{k}-k(\lambda')^{t-k}(1-\lambda')^{k-1}\right).$$

Notice, however, that for any $j = 0, 1, \ldots, k-1$, we may write

$$\sum_{k=0}^{j}\binom{t}{k}\left((t-k)(\lambda')^{t-k-1}(1-\lambda')^{k}-k(\lambda')^{t-k}(1-\lambda')^{k-1}\right)$$
$$=\sum_{k=0}^{j}\binom{t}{k}(t-k)(\lambda')^{t-k-1}(1-\lambda')^{k}-\sum_{k=1}^{j}\binom{t}{k}k(\lambda')^{t-k}(1-\lambda')^{k-1}$$
$$=\sum_{k=0}^{j}\binom{t}{k}(t-k)(\lambda')^{t-k-1}(1-\lambda')^{k}-\sum_{k=0}^{j-1}\binom{t}{k+1}(k+1)(\lambda')^{t-k-1}(1-\lambda')^{k}$$
$$=\binom{t}{j}(t-j)(\lambda')^{t-j-1}(1-\lambda')^{j}+\sum_{k=0}^{j-1}\left(\binom{t}{k}(t-k)-\binom{t}{k+1}(k+1)\right)(\lambda')^{t-k-1}(1-\lambda')^{k}$$
$$=t\binom{t-1}{j}(\lambda')^{t-j-1}(1-\lambda')^{j},$$

where the final equality makes use of the fact that

$$\binom{t}{k}(t-k)-\binom{t}{k+1}(k+1)=\frac{t!(t-j)}{(t-j)!j!}-\frac{t!(j+1)}{(t-j-1)!(j+1)!}=0.$$

Therefore, the first and second lines of the expression for $\partial\hat{U}_0(\lambda',\lambda)/\partial\lambda'$ sum to zero; that is,

$$\frac{\partial\hat{U}_0(\lambda',\lambda)}{\partial\lambda'}=\sum_{t=1}^{T}\delta^{t}\sum_{k=0}^{t}\binom{t}{k}\left(\lambda^{t-k}(1-\lambda)^{k}-(\lambda')^{t-k}(1-\lambda')^{k}\right)\sum_{j=k+1}^{t}\frac{\partial q_t^*(j,\lambda')}{\partial\lambda'}\Delta u^{t-j}d^{j-1}.$$
$$=\sum_{t=1}^{T}\delta^{t}\left(\mathbb{E}[\Phi_t(\kappa_t)|\lambda]-\mathbb{E}[\Phi_t(\kappa_t)|\lambda']\right),$$

where $\Phi_t(k) := \sum_{j=k+1}^{t}\frac{\partial q_t^*(j,\lambda')}{\partial\lambda'}\Delta u^{t-j}d^{j-1}$ is a decreasing function of k and κ_t is a random variable drawn from a binomial distribution with parameters t and $\mu \in \{\lambda, \lambda'\}$. Therefore, the stochastic ordering of binomial distributions implies that, for all $\lambda \in \Lambda$,

$$\frac{\partial\hat{U}_0(\lambda',\lambda)}{\partial\lambda'}\begin{cases}>0 & \text{if } \lambda'<\lambda,\\ =0 & \text{if } \lambda'=\lambda,\\ <0 & \text{if } \lambda'>\lambda.\end{cases}$$

Thus, holding λ fixed, $\hat{U}_0(\lambda',\lambda)$ is maximized when $\lambda' = \lambda$—the buyer has no incentive to misreport her initial-period private information. As established in the main text, the mechanism $(\mathbf{q}^*, \mathbf{p}^*)$ is incentive compatible in all periods $t \geq 1$, and so this mechanism is, indeed, optimal. □

PROOF OF LEMMA 5.1. Fix any $t = 1, \ldots, T$, and notice that the period-t single-deviation constraint (IC-t) may be rewritten as

$$U_t(\alpha_t, \alpha^{t-1}, \lambda) = \max_{\alpha'_t} \Bigg\{ q_t(\alpha'_t, \alpha^{t-1}, \lambda)\alpha_t v(\alpha^{t-1}) - p_t(\alpha'_t, \alpha^{t-1}, \lambda)$$
$$+ \sum_{s=t+1}^{T} \delta^{s-t} \int_{A^{s-t}} \Big(q_s(\alpha^s_{-t}, \alpha'_t, \alpha^{t-1}, \lambda) v(\alpha^s_{-t}, \alpha^{t-1})\alpha_t - p_s(\alpha^s_{-t}, \alpha'_t, \alpha^{t-1}, \lambda) \Big) \, \mathrm{d}W(\alpha^s_{-t}|\lambda) \Bigg\}.$$

Thus, $U_t(\alpha_t, \cdot)$ is an affine maximizer, and therefore a convex function of α_t. Moreover, standard techniques imply that we can rewrite the expressions above as

$$U_t(\alpha_t, \alpha^{t-1}, \lambda) \geq U_t(\alpha'_t, \alpha^{t-1}, \lambda) + \bar{q}_t(\alpha_t, \alpha^{t-1}, \lambda)(\alpha_t - \alpha'_t).$$

Thus, $\bar{q}_t(\alpha_t, \cdot)$ is a subderivative of $U_t(\alpha_t, \cdot)$. But since the $U_t(\alpha_t, \cdot)$ is convex, it is absolutely continuous and, hence, differentiable almost everywhere. Moreover, whenever the partial derivative exists, it must equal its subderivative. Finally, convexity implies that this partial derivative must be a nondecreasing function of α_t. Thus, the period-t single-deviation constraint (IC-t) implies conditions (MON-t) and (IC$'$-t).

In addition, recall that every absolutely continuous function is equal to the definite integral of its derivative. Therefore, for all $\alpha^t \in A^t$ and $\lambda \in \Lambda$,

$$U_t(\alpha_t, \alpha^{t-1}, \lambda) = U_t(\underline{\alpha}, \alpha^{t-1}, \lambda) + \int_{\underline{\alpha}}^{\alpha_t} \bar{q}_t(\alpha'_t, \alpha^{t-1}, \lambda) \, \mathrm{d}\alpha'_t.$$

But since the quantity schedule $q_t(\cdot)$ is non-negative, the integrand above is also non-negative; therefore, condition (IR-t) is satisfied only if $U_t(\underline{\alpha}, \alpha^{t-1}, \lambda) \geq 0$.

Note that the sufficiency of the localized conditions derived above for the period-t single-deviation and participation constraints essentially follows from the Fundamental Theorem of Calculus and monotonicity of the (expected) allocation $\bar{q}_t$. □

PROOF OF LEMMA 5.2. Recall from Equation (17) that $\hat{U}_0(\lambda', \lambda)$ is defined as

$$\hat{U}_0(\lambda', \lambda) := -p_0(\lambda') + \sum_{t=1}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda') v(\alpha^t) - p_t(\alpha^t, \lambda') \right) \, \mathrm{d}W^t(\alpha^t|\lambda).$$

So, note that for any $s \leq T$, we may write

$$\sum_{t=s}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda') v(\alpha^t) - p_t(\alpha^t, \lambda') \right) \, \mathrm{d}W^t(\alpha^t|\lambda)$$
$$= \sum_{t=s}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda') v(\alpha^t) - p_t(\alpha^t, \lambda') \right) \, \mathrm{d}W^t(\alpha^t|\lambda)$$
$$+ \sum_{t=s+1}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda') v(\alpha^t) - p_t(\alpha^t, \lambda') \right) \, \mathrm{d}W^{t-s}(\alpha^t_{-s}|\lambda') \, \mathrm{d}W^s(\alpha^s|\lambda)$$
$$- \sum_{t=s+1}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda') v(\alpha^t) - p_t(\alpha^t, \lambda') \right) \, \mathrm{d}W^{t-s}(\alpha^t_{-s}|\lambda') \, \mathrm{d}W^s(\alpha^s|\lambda).$$

However, it is straightforward to see that, for all $s \leq T$,

$$\delta^s \int_{A^s} \left(q_s(\alpha^s, \lambda')v(\alpha^s) - p_s(\alpha^s, \lambda')\right) \, \mathrm{d}W^s(\alpha^s|\lambda)$$
$$+ \sum_{t=s+1}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda')v(\alpha^t) - p_t(\alpha^t, \lambda')\right) \, \mathrm{d}W^{t-s}(\alpha^t_{-s}|\lambda') \, \mathrm{d}W^s(\alpha^s|\lambda) = \delta^s \int_{A^s} U_s(\alpha^s, \lambda') \, \mathrm{d}W^s(\alpha^s|\lambda)$$

and

$$\sum_{t=s+1}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda')v(\alpha^t) - p_t(\alpha^t, \lambda')\right) \, \mathrm{d}W^{t-s}(\alpha^t_{-s}|\lambda') \, \mathrm{d}W^s(\alpha^s|\lambda)$$
$$= \delta^{s+1} \int_{A^{s+1}} U_{s+1}(\alpha_{s+1}, \alpha^s, \lambda') \, \mathrm{d}G(\alpha_{s+1}|\lambda') \, \mathrm{d}W^s(\alpha^s|\lambda).$$

Therefore, we may write

$$\begin{aligned}\hat{U}_0(\lambda', \lambda) = -p_0(\lambda') &+ \delta \int_A U_1(\alpha_1, \lambda') \, \mathrm{d}G(\alpha_1|\lambda) \\ &- \delta^2 \int_{A^2} U_2(\alpha^2, \lambda') \, \mathrm{d}G(\alpha_2|\lambda') \, \mathrm{d}G(\alpha_1|\lambda) \\ &+ \sum_{t=2}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda')v(\alpha^t) - p_t(\alpha^t, \lambda')\right) \, \mathrm{d}W^t(\alpha^t|\lambda).\end{aligned}$$

Substituting in from the expressions above yields

$$\begin{aligned}\hat{U}_0(\lambda', \lambda) = -p_0(\lambda') &+ \delta \int_A U_1(\alpha_1, \lambda') \, \mathrm{d}G(\alpha_1|\lambda) - \delta^2 \int_{A^2} U_2(\alpha_2, \alpha_2, \lambda') \, \mathrm{d}G(\alpha_2|\lambda') \, \mathrm{d}G(\alpha_1|\lambda) \\ &+ \delta^2 \int_{A^2} U_2(\alpha_2, \alpha_1, \lambda') \, \mathrm{d}G(\alpha_2|\lambda) \, \mathrm{d}G(\alpha_1|\lambda) - \delta^3 \int_{A^3} U_3(\alpha^3, \lambda') \, \mathrm{d}G(\alpha_3|\lambda') \, \mathrm{d}W^2(\alpha^2|\lambda) \\ &+ \sum_{t=3}^{T} \delta^t \int_{A^t} \left(q_t(\alpha^t, \lambda')v(\alpha^t) - p_t(\alpha^t, \lambda')\right) \, \mathrm{d}W^t(\alpha^t|\lambda).\end{aligned}$$

Proceeding inductively in this manner (and making use of the fact that $\mathrm{d}G(\alpha|\mu) = g(\alpha|\mu) \, \mathrm{d}\alpha$ for all $\mu \in \Lambda$), we may conclude that

$$\begin{aligned}\hat{U}_0(\lambda', \lambda) = -p_0(\lambda') &+ \delta \int_A U_1(\alpha_1, \lambda') g(\alpha_1|\lambda) \, \mathrm{d}\alpha_1 \\ &+ \sum_{t=2}^{T} \delta^t \int_{A^t} U_t(\alpha^t, \lambda')(g(\alpha_t|\lambda) - g(\alpha_t|\lambda')) \, \mathrm{d}W^{t-1}(\alpha^{t-1}|\lambda) \, \mathrm{d}\alpha_t.\end{aligned}$$

Furthermore, recall that the constraint (IC-0) requires that $U_0(\lambda) = \max_{\lambda'}\{\hat{U}_0(\lambda', \lambda)\}$ for all $\lambda \in \Lambda$. Therefore, the envelope theorem (see Milgrom and Segal (2002)) implies that

$$\begin{aligned}U_0'(\lambda) = \frac{\partial}{\partial \lambda} \hat{U}_0(\lambda', \lambda)\bigg|_{\lambda'=\lambda} &= \sum_{t=1}^{T} \delta^t \int_{A^t} U_t(\alpha^t, \lambda) \frac{\partial g(\alpha_t|\lambda)}{\partial \lambda} \, \mathrm{d}W^{t-1}(\alpha^{t-1}|\lambda) \, \mathrm{d}\alpha_t \\ &= \sum_{t=1}^{T} \delta^t \int_{A^{t-1}} \left[U_t(\alpha^t, \lambda) \frac{\partial G(\alpha_t|\lambda)}{\partial \lambda}\right]_{\alpha_t = \underline{\alpha}}^{\bar{\alpha}} \, \mathrm{d}W^{t-1}(\alpha^{t-1}|\lambda) \\ &\quad - \sum_{t=1}^{T} \delta^t \int_{A^t} \frac{\partial U_t(\alpha^t, \lambda)}{\partial \alpha_t} \frac{\partial G(\alpha_t|\lambda)}{\partial \lambda} \, \mathrm{d}W^{t-1}(\alpha^{t-1}|\lambda) \, \mathrm{d}\alpha_t,\end{aligned}$$

where the final equality follows from integration by parts. Note, however, that $G(\underline{\alpha}|\lambda) = 0$ for all λ, and $G(\bar{\alpha}|\lambda) = 1$ for all λ; therefore, $\partial G(\underline{\alpha}, \lambda)/\partial\lambda = \partial G(\bar{\alpha}, \lambda)/\partial\lambda = 0$. Substituting in the expression for $\partial U_t/\partial\alpha_t$ from (IC′-t) then yields

$$U_0'(\lambda) = -\sum_{t=1}^{T}\delta^t \int_{A^t} \bar{q}_t(\alpha^t,\lambda)\frac{\partial G(\alpha_t|\lambda)}{\partial\lambda}\,\mathrm{d}W^{t-1}(\alpha^{t-1}|\lambda)\,\mathrm{d}\alpha_t$$
$$= -\sum_{t=1}^{T}\delta^t \int_{A^t} \bar{q}_t(\alpha^t,\lambda)\frac{\partial G(\alpha_t|\lambda)/\partial\lambda}{g(\alpha_t|\lambda)}\,\mathrm{d}W^{t}(\alpha^{t}|\lambda).$$

Finally, note that $q_t(\cdot)$ is non-negative for all t, implying that $\bar{q}_t$ is also non-negative for all t. In addition, $\partial G(\alpha|\lambda)/\partial\lambda \leq 0$ for all $\alpha \in A$ by first-order stochastic dominance. Therefore, $U_0'(\lambda)$ is positive and U_0 is an increasing function. This implies that we can replace the period-zero participation constraint (IR-0) with the requirement that $U_0(\underline{\lambda}) \geq 0$. □

PROOF OF THEOREM 3. Making use of the definition of $\mathbf{p}^*$ from Equations (24) and (25), we may rewrite $\hat{U}_0(\lambda',\lambda)$ from Equation (17) as

$$\hat{U}_0(\lambda',\lambda) = \sum_{t=1}^{T}\delta^t \int_{A^t}\int_{\underline{\alpha}^t}^{v(\alpha^t)} \hat{q}_t^*(v',\lambda')\,\mathrm{d}v'\,\mathrm{d}W^t(\alpha^t|\lambda)$$
$$-\sum_{t=1}^{T}\delta^t \int_{A^t}\int_{\underline{\alpha}^t}^{v(\alpha^t)} \hat{q}_t^*(v',\lambda')\,\mathrm{d}v'\,\mathrm{d}W^t(\alpha^t|\lambda')$$
$$-\sum_{t=1}^{T}\delta^t \int_{\underline{\lambda}}^{\lambda'}\int_{A^t} \bar{q}_t^*(\alpha^t,\mu)\frac{\partial G(\alpha_t|\mu)/\partial\lambda}{g(\alpha_t|\mu)}\,\mathrm{d}W^t(\alpha^t|\mu)\,\mathrm{d}\mu.$$

Taking the partial derivative of this expression with respect to λ' yields

$$\frac{\partial\hat{U}_0(\lambda',\lambda)}{\partial\lambda'} = \sum_{t=1}^{T}\delta^t \int_{A^t}\int_{\underline{\alpha}^t}^{v(\alpha^t)} \frac{\partial\hat{q}_t^*(v',\lambda')}{\partial\lambda}\,\mathrm{d}v'\,\mathrm{d}W^t(\alpha^t|\lambda)$$
$$-\sum_{t=1}^{T}\delta^t \int_{A^t}\int_{\underline{\alpha}^t}^{v(\alpha^t)} \frac{\partial\hat{q}_t^*(v',\lambda')}{\partial\lambda}\,\mathrm{d}v'\,\mathrm{d}W^t(\alpha^t|\lambda')$$
$$-\sum_{t=1}^{T}\delta^t \int_{A^t}\int_{\underline{\alpha}^t}^{v(\alpha^t)} \hat{q}_t^*(v',\lambda')\,\mathrm{d}v'\left(\sum_{s=1}^{t}\frac{\partial g(\alpha_s|\lambda')/\partial\lambda}{g(\alpha_s|\lambda')}\right)\mathrm{d}W^t(\alpha^t|\lambda')$$
$$-\sum_{t=1}^{T}\delta^t \int_{A^t} \bar{q}_t^*(\alpha^t,\lambda')\frac{\partial G(\alpha_t|\lambda')/\partial\lambda}{g(\alpha_t|\lambda')}\,\mathrm{d}W^t(\alpha^t|\lambda').$$

Recall from Equation (21), however, that

$$\sum_{t=1}^{T}\delta^t \int_{A^t} \bar{q}_t(\alpha^t,\lambda')\frac{\partial G(\alpha_t|\lambda')/\partial\lambda}{g(\alpha_t|\lambda')}\,\mathrm{d}W^t(\alpha^t|\lambda')$$
$$= \sum_{t=1}^{T}\delta^t \int_{A^t} q_t(\alpha^t,\lambda')\sum_{s=1}^{t} v(\alpha_{-s}^t,\alpha^{s-1})\frac{\partial G(\alpha_s|\lambda')/\partial\lambda}{g(\alpha_s|\lambda')}\,\mathrm{d}W^t(\alpha^t|\lambda')$$
$$= \sum_{t=1}^{T}\delta^t \sum_{s=1}^{t}\int_{A^{t-1}}\int_{A} \hat{q}_t^*(\alpha_s v(\alpha_{-s}^t,\alpha^{s-1}),\lambda')v(\alpha_{-s}^t,\alpha^{s-1})\frac{\partial G(\alpha_s|\lambda')}{\partial\lambda}\,\mathrm{d}\alpha_s\,\mathrm{d}W^{t-1}(\alpha_{-s}^t,\alpha^{s-1}|\lambda').$$

Straightforward integration by parts implies that

$$\int_A \hat{q}_t^*(\alpha_s v(\alpha_{-s}^t, \alpha^{s-1}), \lambda') v(\alpha_{-s}^t, \alpha^{s-1}) \frac{\partial G(\alpha_s|\lambda')}{\partial \lambda} \, \mathrm{d}\alpha_s = - \int_A \int_{\underline{\alpha}^t}^{\alpha_s v(\alpha_{-s}^t, \alpha^{s-1})} \hat{q}_t^*(v', \lambda') \, \mathrm{d}v' \frac{\partial g(\alpha_s|\lambda')}{\partial \lambda} \, \mathrm{d}\alpha_s,$$

and so

$$\begin{aligned} \sum_{t=1}^T \delta^t \int_{A^t} \bar{q}_t(\alpha^t, \lambda') \frac{\partial G(\alpha_t|\lambda')/\partial \lambda}{g(\alpha_t|\lambda')} \, \mathrm{d}W^t(\alpha^t|\lambda') \\ = - \sum_{t=1}^T \delta^t \sum_{s=1}^t \int_{A^t} \int_{\underline{\alpha}^t}^{v(\alpha^t)} \hat{q}_t^*(v', \lambda') \, \mathrm{d}v' \frac{\partial g(\alpha_s|\lambda')/\partial \lambda}{g(\alpha_s|\lambda')} \, \mathrm{d}W^t(\alpha^t|\lambda'). \end{aligned}$$

Thus, we may conclude that

$$\frac{\partial \hat{U}_0(\lambda', \lambda)}{\partial \lambda'} = \sum_{t=1}^T \delta^t \int_{A^t} \int_{\underline{\alpha}^t}^{v(\alpha^t)} \frac{\partial \hat{q}_t^*(v', \lambda')}{\partial \lambda} \, \mathrm{d}v' \, \mathrm{d} \left[W^t(\alpha^t|\lambda) - W^t(\alpha^t|\lambda') \right].$$

Note, however, that Assumption B.3 implies that, for all $t = 1, \ldots, T$, $\partial \hat{q}_t^*(v', \lambda')/\partial \lambda \geq 0$ for all $v' \in [\underline{\alpha}^t, \bar{\alpha}^t]$ and $\lambda' \in \Lambda$, and therefore

$$\int_{\underline{\alpha}^t}^{v(\alpha^t)} \frac{\partial \hat{q}_t^*(v', \lambda')}{\partial \lambda} \, \mathrm{d}v'$$

is an increasing functions of α_s for all $s = 1, \ldots, t$. The fact that $\{G(\cdot|\lambda)\}_{\lambda \in \Lambda}$ is ordered by first-order stochastic dominance then implies that, for all $\lambda \in \Lambda$,

$$\frac{\partial \hat{U}_0(\lambda', \lambda)}{\partial \lambda'} \begin{cases} > 0 & \text{if } \lambda' < \lambda, \\ = 0 & \text{if } \lambda' = \lambda, \\ < 0 & \text{if } \lambda' > \lambda. \end{cases}$$

Thus, holding λ fixed, $\hat{U}_0(\lambda', \lambda)$ achieves a global maximum when $\lambda' = \lambda$, implying that period-zero single-deviation constraint (IC-0) is satisfied.

Finally, note that (as discussed earlier) Assumptions B.1 and B.2 imply that the buyer is always incentivized to report her private information truthfully in any period $t \geq 1$, regardless of her reports (or misreports) in previous periods. Therefore, the contract $(\mathbf{q}^*, \mathbf{p}^*)$ not only solves the seller's relaxed problem ($\mathcal{Q}'$), but is fully incentive compatible: the buyer prefers truthful reporting to any potential deviation, regardless how complex. □

References

ATHEY, S., AND I. SEGAL (2007a): "Designing Efficient Mechanisms for Dynamic Bilateral Trading Games," *American Economic Review*, 97(2), 131–136.

——— (2007b): "An Efficient Dynamic Mechanism," Unpublished manuscript, Harvard University.

BARON, D. P., AND D. BESANKO (1984): "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, 1(3), 267–302.

BATTAGLINI, M. (2005): "Long-Term Contracting with Markovian Consumers," *American Economic Review*, 95(3), 637–658.

BERGEMANN, D., AND M. SAID (2011): "Dynamic Auctions," in *Wiley Encyclopedia of Operations Research and Management Science*, ed. by J. J. Cochran, pp. 1511–1522. Wiley: New York.

BERGEMANN, D., AND J. VÄLIMÄKI (2010): "The Dynamic Pivot Mechanism," *Econometrica*, 78(2), 771–789.

BESANKO, D. (1985): "Multi-Period Contracts Between Principal and Agent with Adverse Selection," *Economics Letters*, 17(1–2), 33–37.

BOARD, S., AND A. SKRZYPACZ (2010): "Revenue Management with Forward-Looking Buyers," Unpublished manuscript, Stanford University.

COURTY, P., AND H. LI (2000): "Sequential Screening," *Review of Economic Studies*, 67(4), 697–717.

COX, J. C., S. A. ROSS, AND M. RUBINSTEIN (1979): "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7(3), 229–263.

DAI, C., T. R. LEWIS, AND G. LOPOMO (2006): "Delegating Management to Experts," *RAND Journal of Economics*, 37(3), 503–520.

DEB, R. (2009): "Optimal Contracting of New Experience Goods," Unpublished manuscript, University of Toronto.

——— (2011): "Intertemporal Price Discrimination with Stochastic Values," Unpublished manuscript, University of Toronto.

ESÖ, P., AND B. SZENTES (2007): "Optimal Information Disclosure in Auctions and the Handicap Auction," *Review of Economic Studies*, 74(3), 705–731.

GERSHKOV, A., AND B. MOLDOVANU (2009a): "Dynamic Revenue Maximization with Heterogeneous Objects: A Mechanism Design Approach," *American Economic Journal: Microeconomics*, 1(2), 168–198.

——— (2009b): "Learning About the Future and Dynamic Efficiency," *American Economic Review*, 99(4), 1576–1587.

——— (2010a): "Demand Uncertainty and Dynamic Allocation with Strategic Agents," Unpublished manuscript, University of Bonn.

——— (2010b): "Efficient Sequential Assignment with Incomplete Information," *Games and Economic Behavior*, 68(1), 144–154.

HARRIS, M., AND A. RAVIV (1981): "A Theory of Monopoly Pricing Schemes with Demand Uncertainty," *American Economic Review*, 71(3), 347–365.

HENDEL, I., AND A. LIZZERI (2003): "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance," *Quarterly Journal of Economics*, 118(1), 299–328.

KAKADE, S. M., I. LOBEL, AND H. NAZERZADEH (2011): "Optimal Dynamic Mechanism Design and the Virtual Pivot Mechanism," Unpublished manuscript, Microsoft Research New England.

KENNAN, J. (2001): "Repeated Bargaining with Persistent Private Information," *Review of Economic Studies*, 68(4), 719–755.

KRÄHMER, D., AND R. STRAUSZ (2011): "Optimal Procurement Contracts with Pre-Project Planning," *Review of Economic Studies*, forthcoming.

KURIBKO, N., AND T. R. LEWIS (2010): "Assigning Property Rights to Ensure Efficient Exchange and Investment," Unpublished manuscript, Duke University.

LOGINOVA, O., AND C. R. TAYLOR (2008): "Price Experimentation with Strategic Buyers," *Review of Economic Design*, 12(3), 165–187.

MATTHEWS, S., AND J. MOORE (1987): "Monopoly Provision of Quality and Warranties: An Exploration in the Theory of Multidimensional Screening," *Econometrica*, 55(2), 441–467.

MILGROM, P., AND I. SEGAL (2002): "Envelope Theorem for Arbitrary Choice Sets," *Econometrica*, 70(2), 583–601.

MIRAVETE, E. J. (2003): "Choosing the Wrong Calling Plan? Ignorance and Learning," *American Economic Review*, 93(1), 297–310.

MUSSA, M., AND S. ROSEN (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18(2), 301–317.

MYERSON, R. B. (1986): "Multistage Games with Communication," *Econometrica*, 54(2), 323–358.

PAI, M. M., AND R. VOHRA (2011): "Optimal Dynamic Auctions and Simple Index Rules," Unpublished manuscript, University of Pennsylvania.

PAVAN, A. (2007): "Long-Term Contracting in a Changing World," Unpublished manuscript, Northwestern University.

PAVAN, A., I. SEGAL, AND J. TOIKKA (2009): "Dynamic Mechanism Design: Revenue Equivalence, Profit Maximization and Information Disclosure," Unpublished manuscript, Northwestern University.

——— (2010): "Infinite-Horizon Mechanism Design: The Independent-Shock Approach," Unpublished manuscript, Northwestern University.

SAID, M. (2011): "Auctions with Dynamic Populations: Efficiency and Revenue Maximization," Unpublished manuscript, Washington University in St. Louis.

VULCANO, G., G. VAN RYZIN, AND C. MAGLARAS (2002): "Optimal Dynamic Auctions for Revenue Management," *Management Science*, 48(11), 1388–1407.

WILSON, R. B. (1993): "Priority Pricing," in *Nonlinear Pricing*, pp. 236–258. Oxford University Press: New York.