Review of Economic Studies (2013) **80**, 1–34 doi:10.1093/restud/rds021

Advance access publication 18 May 2012

Progressive Screening: Long-Term Contracting with a Privately Known Stochastic Process

RAPHAEL BOLESLAVSKY

Department of Economics, University of Miami

and

MAHER SAID

Olin Business School, Washington University in St. Louis

First version received July 2009; final version accepted April 2012 (Eds.)

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, the realized values are also private information. We characterize a class of environments in which the profit-maximizing long-term contract offered by a monopolist takes an especially simple structure: we derive sufficient conditions on primitives under which the 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 inefficiently 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 process information in the initial period and then progressively screens across realized values so as to reduce the information rents paid in future periods.

Key words: Asymmetric information, Dynamic incentives, Dynamic mechanism design, Long-term contracts, Sequential screening

JEL Codes: C73, D82, D86

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, but instead may be changing over time. 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; an employer need not observe the changes in an employee's productivity; and a downstream retailer need not know the effectiveness of an upstream manufacturer's investments in cost reduction. Clearly, optimal long-term contracts must be designed in order to account for these dynamic information asymmetries.

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. Returning to the examples above, only the buyer knows how many complementary products she plans to buy; only the employee knows the likelihood of distractions arising at home that affect her productivity; and only the manufacturer knows its ability to implement process innovations.

We analyze these issues 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 that governs the evolution of her value.

Formally, we set out to characterize the profit-maximizing T-period contract (where T is potentially infinite) for a single seller facing a buyer whose value evolves according to a privately known stochastic process. In the initial period, the buyer privately observes a parameter λ. In each subsequent period, she privately observes a random shock α, and her value is the product of all previous shocks.[1] The conditional distributions of shocks are ranked by first-order stochastic dominance, so that a buyer with a higher value of λ is more likely to experience "good" shocks and have higher values in each period. We assume that the seller has the ability to fully commit to arbitrary long-term contractual forms.[2] Therefore, 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 in every period.

Our main result is a characterization of a class of environments in which the optimal long-term contract takes an especially simple structure. More specifically, we find simple sufficient conditions on the distribution of λ and the conditional distributions of shocks under which incentives can be decoupled over time. We show that, when this is the case, the optimal dynamic contract is a menu of deterministic sequences of static contracts that progressively screen the buyer's values: the seller introduces additional supply restrictions over time and eventually excludes all types in order to extract rents from higher-valued buyers.

In our baseline model, we assume that the buyer has single-unit demand in each period, and that 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 $1-\lambda$. In the benchmark case where the good is produced in each period at zero cost, the seller commits to a finite menu of price plans, each of which presents the buyer with an entry fee and a predetermined sequence of future prices for the good. 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 that period's prespecified price.[3] Since 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. (These qualitative features of the optimal contract carry over to the case of positive marginal cost.)

Moreover, each of these price plans begins with a finite-length "honeymoon" phase. In each period of this phase, the price for the good grows by a factor d; after the honeymoon phase ends, the price grows in each period by the larger factor u. This implies that the buyer purchases the good throughout the honeymoon phase, regardless of the realized shocks to her value. After this phase,

1. Note that new information arrives exogenously, unlike the endogenous information acquisition in Bergemann and Välimäki (2002); Gershkov and Szentes (2009); Krähmer and Strausz (2011); and Shi (2012), among others.

2. 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.

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

however, the price grows deterministically, while the buyer's value grows only stochastically—with a sufficiently long time horizon, the seller terminates the relationship inefficiently early. Since prices in a plan with a longer honeymoon phase are always lower than those in plans with a shorter honeymoon phase, longer honeymoons are attractive to all buyers. However, the entry fees for the various plans are increasing in the length of their honeymoon phases. In order to justify paying a larger initial fee, the buyer must therefore anticipate that her future values will be sufficiently high that the lower future prices fully compensate for the initial fee—paying a larger entry fee is justified only if the probability of good shocks λ is sufficiently high. Thus, the various entry fees and honeymoon phase length serves to screen across realizations of λ, whereas the post-honeymoon-phase growth in prices serve to restrict supply to lower-valued buyers, reducing the rents paid to higher-valued buyers.

We also extend our analysis to the setting where the seller faces an increasing convex cost function and drop the single-unit demand assumption. We also assume that the buyer's valuation shocks are independently drawn from a family of continuous distributions parametrized by λ and ordered by first-order stochastic dominance, so that larger realizations of λ generate higher values.

In this more general environment, we derive sufficient conditions on the underlying primitives under which incentives decouple over time; thus, the optimal contract again consists of a sequence of static contracts. In the initial period, the buyer chooses (on the basis of λ) from a continuum of contingent price–quantity schedules, each of which is a fixed sequence of price–quantity menus that screen across future values. As is standard in nonlinear pricing problems, each of these menus provides greater quantities to buyers that report higher shocks. Moreover, these menus feature more generous quantity provision for buyers reporting higher values of λ, excluding fewer realized valuations and allocating larger quantities to included buyers. Within a given sequence of menus, however, the quantity schedules become less permissive over time as the seller "tightens the screws": the set of period-$(t+1)$ reports that prevent exclusion is a subset of the corresponding period-t set of reports. Thus, as in the discrete-shock case, the seller inefficiently restricts supply in order to extract additional rents, with greater restrictions for buyers that report lower values of λ. Similarly, the prices within each period's menu in this optimal contract 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 initial private information with entry fees and the generosity of future menus, and then progressively screens across realized values with nonlinear prices in future periods.

As is typically the case in dynamic mechanism design, the primary hurdle we face in solving the seller's problem 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 multistage deviations from truthfulness. This generates a complex and relatively intractable set of constraints that must be satisfied by any optimal contract. One common approach in the literature for dealing with this issue is to restrict attention to two-period models—this is the approach of, among others, Baron and Besanko (1984); Courty and Li (2000); Esö and Szentes (2007); Krähmer and Strausz (2011); and Riordan and Sappington (1987a,b). In such models, it is possible to simplify the second-period constraints and work backward to the first period; this methodology does not, however, generalize easily to the longer time horizons we study in the present work.

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. Specifically, we impose a set of single-deviation constraints that rule out "one-time" deviations from truthful reporting: in each period t, the buyer must prefer truthful reporting to any possible misreport, assuming that *all* future shocks are reported truthfully.[4] We then provide easily verified sufficient conditions on the underlying environment under which the solution to this relaxed problem depends only on the buyer's initial type and realized values, but not on the particular sequence of shocks generating those values. By pairing this allocation rule with a payment scheme that is also path-independent, we decouple the buyer's incentives in any one period from those in the next. This guarantees that truthtelling is an optimal continuation strategy for the buyer, regardless of her history of past reports or misreports. This property implies that the restricted class of constraints in our relaxed problem is, in fact, sufficient for "global" incentive compatibility, thereby justifying our approach.

Our paper contributes to the growing literature on *optimal* dynamic mechanism design that focuses on the design of profit-maximizing mechanisms in dynamic settings.[5] While much of the recent work in this area focuses on settings where agents arrive and depart dynamically over time while their private information remains fixed, our paper joins another strand of the literature where the population of agents is fixed, but their private information changes over time.[6]

Baron and Besanko (1984) were the first to study dynamic contracting with changing types, deriving necessary conditions for optimality in a two-period model using an "informativeness measure" of initial-period private information on future types. Courty and Li (2000) study a two-period model where consumers are initially uncertain about their future demand but receive additional private information before consumption.[7] In contrast, our focus in the present work is on arbitrarily long time horizons. 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 that cannot arise in a two-period model.[8] In addition, the longer time horizon necessitates consideration of a richer set of incentive compatibility constraints, as the buyer may misreport her value multiple times in an attempt to take advantage of future contractual terms. As discussed above, such compound deviations introduce additional technical difficulties in identifying the optimal contract that preclude the use of backward induction common in two-period models.

Besanko (1985) and Battaglini (2005) also explore optimal contracting in dynamic settings with more than two periods. In Besanko's model, the buyer's values follow a first-order autoregressive process where each period's value is a linear function of the previous value and an i.i.d. shock. As in our model, the buyer's initial-period type exerts a persistent influence on

4. This approach was first used in a two-period problem by Esö and Szentes (2007) and extended to longer time horizons by Pavan, Segal and Toikka (2011).

5. There is also a parallel literature focusing on *efficient* dynamic mechanism design; see, among others, Athey and Segal (2007a,b); Bergemann and Välimäki (2010); Gershkov and Moldovanu (2009b, 2010a,b); and Kuribko and Lewis (2010). See Bergemann and Said (2011) for a survey of both literatures.

6. See, among others, Board and Skrzypacz (2010); Gershkov and Moldovanu (2009a); Mierendorff (2011); Pai and Vohra (2011); Said (2012); and Vulcano, van Ryzin, and Maglaras (2002) for models with a dynamic population and static types, and Board (2007) and Deb (2009, 2011) for recent examples with a fixed population and changing types. Garrett (2011) is a recent contribution that combines both dynamic arrivals and changing private information.

7. Dai, Lewis, and Lopomo (2006) and Riordan and Sappington (1987b) examine similar issues in a procurement setting, while Esö and Szentes (2007) and Riordan and Sappington (1987a) consider a two-period setting in which the principal faces multiple competing agents. Meanwhile, Miravete (2003) empirically demonstrates the importance of sequential screening considerations in the design of contracts for telephone service.

8. Since many allocations are made over time, we also find it more compelling to consider contracts where the buyer pays for consumption in each period instead of refund contracts as in Courty and Li (2000). It is straightforward, however, to show that the payment rule may be modified to make use of the refunds instead of prices without affecting incentives—as is standard in dynamic incentive problems, there can be many payment rules supporting the same allocation rule.

all future values; this generates decreasing distortions if the process is stationary and increasing distortions when it is not. Battaglini, on the other hand, studies a model where the buyer's value evolves according to a two-state Markov process with commonly known transition probabilities, so the shock in period t depends directly on the realized value in period $t-1$. In that setting, the distribution of future values converges to a steady-state distribution, so the impact of the buyer's initial type decreases over time and the optimal contract is asymptotically efficient. In the present work, however, shocks are *conditionally* independent (given λ); therefore, each shock induces greater dependence of values on the buyer's initial type. This increasing dependence is the source of the increasing distortions and inefficiency in our environment's optimal contract.

Our use of a relaxed problem that imposes only single-deviation incentive constraints to circumvent the difficulties of compound deviations and dynamic incentive compatibility resembles the approaches of Esö and Szentes (2007) and Pavan, Segal and Toikka (2011). Esö and Szentes observe that any stochastic process governing values may be transformed into a sequence of independent shocks. This transformation transfers the dependence of value shocks into more complex payoff functions; in their two-period model, however, they are able to provide sufficient conditions for implementation of the optimal allocation. Pavan, Segal and Toikka use a similar observation to derive a dynamic envelope formula for arbitrary time horizons and stochastic processes.[9] This dynamic envelope formula is used to extend the standard static payoff equivalence result to dynamic settings, and then to identify sufficient conditions for incentive compatibility. While our continuous–discrete setup with conditionally independent shocks requires different arguments to derive the optimal contract, their unifying framework helps explain how distortions depend on the "impulse response" of future payoffs to private information at the time of contracting. In particular, a mechanism designer distorts decision in order to account for the buyer's informational advantage at the time of contracting, and these distortions are most effective at histories where the buyer's values are most responsive to her initial type. Since each additional shock in our model compounds the dependence of values on λ, the induced value distributions in later periods are *more* sensitive to the initial type than those in earlier periods. This results in progressive screening and increasingly aggressive exclusion of buyers over the course of the relationship.

2. ENVIRONMENT

We consider a dynamic setting in which a buyer repeatedly purchases a nondurable good from a single seller. When the buyer pays a price p and receives quantity q of the good in period t, 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 stochastic sequence 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. We will denote α^t by 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

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

9. Kakade, Lobel and Nazerzadeh (2011) also use an independent shock representation, but their approach imposes an additional "separability" assumption and requires the agent to report her entire private history in each period.

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

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

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 $G(\cdot|\lambda)$ with support $A \subseteq \mathbb{R}_+$. Moreover, we assume that the family $\{G(\cdot|\lambda)\}$ is ordered in terms of first-order stochastic dominance; that is, $G(\cdot|\lambda)$ first-order stochastically dominates $G(\cdot|\lambda')$ whenever $\lambda > \lambda'$.

At the time of contracting (which we take to be 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 drawn from the distribution F on an interval $\Lambda \subseteq \mathbb{R}_+$. We assume that f, the density of F, is strictly positive and differentiable on Λ.

In each period $t \geq 1$, the seller can produce q units of the good at a cost $c(q)$. The relationship between the buyer and the seller persists for $T \leq \infty$ periods, and is discounted with the common discount factor $\delta \in (0,1]$. (If $T = \infty$, we require the additional restriction that $\delta \int_A \alpha \, \mathrm{d}G(\alpha|\lambda) < 1$ for all $\lambda \in \Lambda$ to guarantee the boundedness of expected payoffs.) In the initial period, the seller offers a long-term contract to the buyer. If the buyer accepts this offer, 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, commitment is one-sided in our model: the buyer is free to break off the relationship at any time.

3. THE SELLER'S PROBLEM

The seller wishes to design and offer a contract that maximizes her expected profits. Since the Myerson (1986) revelation principle for multistage games holds in our environment, 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.

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 agent's 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 $\widehat{H}_t$; the buyer's strategy, given the seller's mechanism, is then simply a sequence of mappings $\hat{r}_t : \widehat{H}_t \to A$ for $t \geq 1$, and $\hat{r}_0 : \Lambda \to \Lambda$.

A direct mechanism is *incentive compatible* if it induces truthful reporting in every period: on the equilibrium path, the agent has no incentive to misreport her new private information. This requires the agent to prefer revealing her private information truthfully to any misreport followed by optimal continuation reporting (which may involve additional misreports). Thus, the complete set of incentive compatibility constraints in our setting is large and potentially intractable.

In order to avoid this complexity, we use an indirect approach to solve 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 provide sufficient conditions guaranteeing that this restricted set of constraints is, in fact, sufficient for "full" incentive compatibility in our setting. 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 when our sufficient conditions are satisfied. 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-zero) 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 $\lambda \in \Lambda$,

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

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

$$\widehat{U}_0(\lambda',\lambda) := -p_0(\lambda') + \sum_{t=1}^{T} \delta^t \int_{A^t} \big(q_t(\alpha^t,\lambda')v(\alpha^t) - p_t(\alpha^t,\lambda')\big)\,\mathrm{d}W^t(\alpha^t|\lambda). \tag{2}$$

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

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

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

$$\begin{aligned} U_t(\alpha^t,\lambda) &:= q_t(\alpha^t,\lambda)v(\alpha^t) - p_t(\alpha^t,\lambda) \\ &+ \sum_{s=t+1}^{T} \delta^{s-t} \int_{A^{s-t}} \big(q_s(\alpha^s_{-t},\alpha^t,\lambda)v(\alpha^s_{-t},\alpha^t) - p_s(\alpha^s_{-t},\alpha^t,\lambda)\big)\,\mathrm{d}W^{s-t}(\alpha^s_{-t}|\lambda). \end{aligned} \tag{3}$$

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

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

Notice that condition (IC-t) is essentially the static incentive compatibility constraint faced by a buyer with private information about α_t alone. In a standard static contracting problem, quasilinearity and a single-crossing condition imply that the incentive compatibility constraints are equivalent to (a) the monotonicity of the allocation rule; and (b) the determination of a buyer's utility (up to a constant) by that allocation rule alone. The buyer in our setting is forward-looking, however, and her utility depends upon her expectations about the future. Naturally, this implies that the "localized" period-t constraints in our relaxed problem will involve the expected discounted value of current and future allocations, which we denote by

$$\bar{q}_t(\alpha^t,\lambda):=q_t(\alpha^t,\lambda)v(\alpha^{t-1})+\sum_{s=t+1}^{T}\delta^{s-t}\int_{A^{s-t}}q_s(\alpha^s_{-t},\alpha^t,\lambda)v(\alpha^s_{-t},\alpha^{t-1})\mathrm{d}W^{s-t}(\alpha^{s-t}|\lambda). \tag{4}$$

Finally, 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 as

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

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

The seller's profit from any feasible contract is then simply 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

$$\begin{aligned}\Pi(\lambda)&:=p_0(\lambda)+\sum_{t=1}^{T}\delta^t\int_{A^t}\big(p_t(\alpha^t,\lambda)-c(q_t(\alpha^t,\lambda))\big)\,\mathrm{d}W^t(\alpha^t|\lambda)\\&=-U_0(\lambda)+\sum_{t=1}^{T}\delta^t\int_{A^t}\big(q_t(\alpha^t,\lambda)v(\alpha^t)-c(q_t(\alpha^t,\lambda))\big)\,\mathrm{d}W^t(\alpha^t|\lambda)\end{aligned} \tag{5}$$

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:

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

4. DISCRETE SHOCKS

We begin by specializing to the setting in which there are only two possible shocks and the buyer's value evolves according to a recombinant binomial tree process with "upward" transition probability λ. In particular, we let $\Lambda:=[0,1]$ and assume that each shock α_t is drawn from the discrete distribution $G(\cdot|\lambda)$ on $A:=\{u,d\}$, where

$$G(\alpha|\lambda)=\lambda H_u(\alpha)+(1-\lambda)H_d(\alpha).$$

($H_z(\cdot)$ is the Heaviside step function centered at $z\in\mathbb{R}$.) We assume that $u>d>0$, and let $\Delta:=u-d$. Thus, the buyer experiences either a "good" shock (u) or a "bad" shock (d) in each period, and the probability λ of experiencing the higher shock is fixed across time.

4.1. *Simplifying the seller's relaxed problem*

We approach the seller's optimal contracting problem by first simplifying the single deviation and participation constraints in the relaxed problem ($\mathcal{R}$). Since (IC-t) is essentially a static incentive compatibility constraint, we have the following "standard" result (whose proof may be found in the Appendix) that the period-t constraints may be replaced by a monotonicity condition and a downward incentive compatibility constraint:

Lemma 4.1. *The period-t incentive compatibility and individual rationality constraints (IC-*t*) and (IR-*t*), where $t=1,\ldots,T$, are satisfied if, and only if, for all $\alpha^{t-1}\in A^{t-1}$ and all $\lambda\in\Lambda$,*

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

$$\bar{q}_t(u,\alpha^{t-1},\lambda)\geq \bar{q}_t(d,\alpha^{t-1},\lambda); \textit{ and} \tag{MON$'$-t}$$

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

Notice that at the time of initial contracting (unlike in period $t\geq 1$), the buyer's private information does not directly affect her flow payoffs. Rather, the realization of λ only affects the buyer's beliefs about the evolution of her future preferences. Therefore, the buyer in period zero 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 the 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)—allows a reformulation of the seller's relaxed problem ($\mathcal{R}$) into one involving only allocation rules (and not the payment rules).

Lemma 4.2. *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*

$$U_0'(\lambda)=\sum_{t=1}^{T}\delta^t\int_{A^{t-1}}\Big(U_t(u,\alpha^{t-1},\lambda)-U_t(d,\alpha^{t-1},\lambda)\Big)\,\mathrm{d}W^{t-1}(\alpha^{t-1}|\lambda). \tag{IC$'$-0}$$

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. In addition, note that the nonnegativity of allocation probabilities implies that $U_0'(\lambda)\geq 0$ whenever condition (IC$'$-t) is satisfied for all t. Therefore, U_0 is increasing in any solution to the seller's problem, and the period-zero participation constraint (IR-0) becomes

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

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

$$\max_{\{\mathbf{q},\mathbf{p}\}} \left\{ \begin{array}{l} -U_0(0)+\displaystyle\int_\Lambda \sum_{t=1}^{T} \delta^t \int_{A^t} \big(q_t(\alpha^t,\lambda)v(\alpha^t)-c(q_t(\alpha^t,\lambda))\big)\,\mathrm{d}W^t(\alpha^t|\lambda)\mathrm{d}F(\lambda) \\ \qquad -\displaystyle\int_\Lambda \sum_{t=1}^{T} \delta^t \int_{A^t} \Big(U_t(u,\alpha^{t-1},\lambda)-U_t(d,\alpha^{t-1},\lambda)\Big)\frac{1-F(\lambda)}{f(\lambda)}\,\mathrm{d}W^t(\alpha^t|\lambda)\mathrm{d}F(\lambda) \end{array} \right\}$$

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

Clearly, $U_0(0)=0$ in any solution to this problem, 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 generating incentives for truthtelling. In addition, it is clear that, for all $(\alpha^{t-1},\lambda)\in A^{t-1}\times\Lambda$, we must minimize $U_t(u,\alpha^{t-1},\lambda)-U_t(d,\alpha^{t-1},\lambda)$. However, the constraints (IC$'$-t) provide a lower bound on this difference, and so these downward incentive constraints must bind. We can then rewrite the seller's objective function as

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

Finally, note that

$$\begin{aligned} &\sum_{t=1}^{T}\delta^t\int_{A^t}\bar{q}_t(d,\alpha^{t-1},\lambda)\mathrm{d}W^t(\alpha^t|\lambda) \\ &\quad=\sum_{t=1}^{T}\delta^t\int_{A^t}\left(q_t(d,\alpha^{t-1},\lambda)v(\alpha^{t-1})\right. \\ &\qquad\left.+\sum_{s=t+1}^{T}\delta^{s-t}\int_{A^{s-t}}q_s(\alpha^s_{-t},d,\alpha^{t-1},\lambda)v(\alpha^s_{-t},\alpha^{t-1})\mathrm{d}W^{s-t}(\alpha^s_{-t}|\lambda)\right)\mathrm{d}W^t(\alpha^t|\lambda) \\ &\quad=\sum_{t=1}^{T}\delta^t\int_{A^t}\left(\sum_{s=t}^{T}\delta^{s-t}\int_{A^{s-t}}q_s(\alpha^s_{-t},d,\alpha^{t-1},\lambda)v(\alpha^s_{-t},\alpha^{t-1})\mathrm{d}W^{s-t}(\alpha^s_{-t}|\lambda)\right)\mathrm{d}W^t(\alpha^t|\lambda) \end{aligned}$$

Interchanging the order of summations, we may write this expression as

$$\sum_{t=1}^{T}\delta^t\int_{A^t}\bar{q}_t(d,\alpha^{t-1},\lambda)\mathrm{d}W^t(\alpha^t|\lambda)=\sum_{t=1}^{T}\sum_{s=1}^{t}\delta^t\int_{A^t}q_t(\alpha^t_{-s},d,\alpha^{s-1},\lambda)v(\alpha^t_{-s},\alpha^{s-1})W^t(\alpha^t|\lambda). \quad (6)$$

Substituting from Equation (6) into the seller's problem then yields the following relaxed problem:

$$\max_{\{\mathbf{q},\mathbf{p}\}} \left\{ \begin{aligned} &\sum_{t=1}^{T} \delta^t \int_\Lambda \int_{A^t} \Bigg(q_t(\alpha^t,\lambda)v(\alpha^t) - c(q_t(\alpha^t,\lambda)) \\ &\qquad - \sum_{s=1}^{t} q_t(\alpha^t_{-s}, d, \alpha^{s-1}, \lambda) v(\alpha^t_{-s}, \alpha^{s-1}) \Delta \frac{1-F(\lambda)}{f(\lambda)} \Bigg) \mathrm{d}W^t(\alpha^t|\lambda)\mathrm{d}F(\lambda) \end{aligned} \right\} \qquad (\mathcal{R}')$$

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

Before proceeding to the solution of the seller's problem, it is helpful to interpret the objective function in $(\mathcal{R}')$, especially by way of comparison with a standard (static) nonlinear pricing setting. For all t and each $\lambda \in \Lambda$, we can rewrite the inner integrand in this objective function as

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

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

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

As in the static mechanism design setting, the first term in this expression is the buyer's contribution to the social surplus, whereas 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. The final term in the expression above reflects the persistent impact of the buyer's initial-period type and future values: as is well-established in the dynamic mechanism design literature, distortions in the optimal contract depend upon the sensitivity of future values to the buyer's initial private information.

To more clearly see that $v(\alpha^t)\sum_{s=1}^{t} \mathbb{1}_d(\alpha_s)[(\Delta/d)/(1-\lambda)]$ is the measure, in our setting, of the informational linkage between λ and v_t, we can use the "independent shock approach" of Esö and Szentes (2007) to characterize the responsiveness of values to changes in λ. In particular, let $\xi_s \in [0,1]$ be a uniform random variable that is independent of λ. We can then write the period-s shock α_s as

$$\tilde{\alpha}_s(\xi_s,\lambda) = d + \Delta H_{1-\lambda}(\xi_s),$$

where $H_{1-\lambda}(\cdot)$ is the Heaviside step function centered at $1-\lambda$. Thus, we can identify the buyer with shock $\alpha_s = d$ with the "average buyer" with $\xi_s < 1-\lambda$, and the buyer with shock $\alpha_s = u$ with the "average buyer" with $\xi_s \geq 1-\lambda$. It is then straightforward to see that

$$\mathbb{E}\left[\frac{\partial \tilde{\alpha}_s(\xi_s,\lambda)}{\partial \lambda}\middle|\tilde{\alpha}_s(\xi_s,\lambda)=u\right]=\Delta\frac{\int_{1-\lambda}^{1}\mathrm{d}H_{1-\lambda}(\xi_s)}{\int_{1-\lambda}^{1}\mathrm{d}\xi_s}=0 \text{ and}$$

$$\mathbb{E}\left[\frac{\partial \tilde{\alpha}_s(\xi_s,\lambda)}{\partial \lambda}\middle|\tilde{\alpha}_s(\xi_s,\lambda)=d\right]=\Delta\frac{\int_{0}^{1-\lambda}\mathrm{d}H_{1-\lambda}(\xi_s)}{\int_{0}^{1-\lambda}\mathrm{d}\xi_s}=\frac{\Delta}{1-\lambda}.$$

Thus, only the "bad" d shocks are responsive to changes in λ. Of course, since the buyer's value is the product of multiple shocks, the overall responsiveness of the period-t value to λ is then

$$\sum_{s=1}^{t} v(\alpha_{-s}^{t},\alpha^{s-1})\mathbb{1}_d(\alpha_s)\frac{\Delta}{1-\lambda}=v(\alpha^t)\sum_{s=1}^{t}\mathbb{1}_d(\alpha_s)\frac{\Delta/d}{1-\lambda}.$$

Thus, in order to guarantee the satisfaction of downward incentive constraints and minimize the information rents paid to buyers with high values, the seller must introduce additional distortions for each reported low shock, where the size of these distortions depend upon λ.

Before moving on, we introduce the following condition on the distribution F of the buyer's initial-period private information:

Condition A. *The distribution F of initial-period private information is such that*

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

is decreasing in λ.

This is a sufficient condition for the monotonicity of the buyer's virtual value: it guarantees that $\varphi(\alpha^t,\lambda)$ is nondecreasing in λ for all $t=1,\ldots,T$ and all $\alpha^t \in A^t$. We should note that this condition is strictly stronger than the standard assumption that F is log-concave; however, similar sufficient conditions are frequently needed in dynamic and multidimensional mechanism design settings.[10] Moreover, this condition is satisfied by a large variety of distributions F on the unit interval. For instance, the uniform distribution, as well as any power distribution $F(\lambda)=\lambda^x$, where $x\geq 1$, satisfies this condition. Similarly, the beta and Kumaraswamy distributions satisfy Condition A whenever their shape parameters are $a\geq 1$ and $b>0$.

4.2. *Single-unit demand and constant marginal cost*

We now solve for the optimal long-term contract for the benchmark case in which the buyer has single-unit demand in each period, and the good is produced at a constant marginal cost, so that $c(q)=cq$ for some constant $c\geq 0$. In this case, the seller's relaxed optimization problem is

$$\max_{\{\mathbf{q},\mathbf{p}\}}\left\{\sum_{t=1}^{T}\delta^t\iint_{\Lambda\times A^t}\left(\varphi(\alpha^t,\lambda)-c)q_t(\alpha^t,\lambda)\right)\mathrm{d}W^t(\alpha^t|\lambda)\mathrm{d}F(\lambda)\right\}$$

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

10. The utility of such conditions was first noted by Baron and Besanko (1984) and Besanko (1985), and an analogous "attribute ordering" condition was imposed by Matthews and Moore (1987) in a multidimensional screening setting.

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

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

Since $[(\Delta/d)/(1-\lambda)][(1-F(\lambda))/f(\lambda)]>0$ for all $\lambda<1$, 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.[11] 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 nonzero when $\lambda=1$, l'Hôpital's rule implies that

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

thus, for any t, a buyer who has experienced more than $1/\gamma$ downward shocks d will have a *negative* virtual value. Clearly, if our sufficient condition on the distribution F is satisfied, then the cutoff $k_t(\lambda)$ is nondecreasing in λ for all t.

Finally, let us denote the optimal allocations in the relaxed problem as

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

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 $\lambda\in\Lambda$. 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 $\alpha_{-t}^s\in A^{s-t}$. 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.

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 that only 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. 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.

11. 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 is never distorted away from the efficient allocation. We do not focus on this, however, as this is a zero probability type.

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 can use Lemma 4.2 (combined with the fact that the constraints (IC$'$-t) and (IR$'$-0) bind) to show that the initial payment must be

$$\begin{aligned} p_0^*(\lambda) :=& \sum_{t=1}^T \delta^t \int_{A^t} \big(q_t^*(\alpha^t,\lambda)v(\alpha^t) - p_t^*(\alpha^t,\lambda)\big)\,\mathrm{d}W^t(\alpha^t|\lambda) - U_0(\lambda) \\ =& \sum_{t=1}^T \delta^t \int_{A^t} \big(q_t^*(\alpha^t,\lambda)v(\alpha^t) - p_t^*(\alpha^t,\lambda)\big)\,\mathrm{d}W^t(\alpha^t|\lambda) \\ & - \sum_{t=1}^T \delta^t \int_0^\lambda \int_{A^t} \bar{q}_t^*(d,\alpha^{t-1},\mu)\Delta\,\mathrm{d}W^t(\alpha^t|\mu)\,\mathrm{d}\mu. \end{aligned} \tag{12}$$

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 nonnegative. 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: overreporting the number of d shocks leads to the exclusion of the buyer in situations where truthful reporting may have led to a profitable allocation, whereas underreporting 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 4.2 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, Condition A is sufficient for incentive compatibility in the initial period. The theorem below (whose proof is found in the Appendix) demonstrates this fact.

Theorem 4.1. *Suppose that the distribution F satisfies Condition 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* $(\mathcal{R}')$.

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 Condition 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 by a factor u in *every* period thereafter. This initial period of slower price growth is essentially a "honeymoon phase" after which the slope of the price path rises. 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. Indeed, in order to justify paying a larger entry fee, the buyer must anticipate that her future surplus will be sufficiently high to fully compensate her for the initial cost—paying a larger initial fee for a future price discount is justified only if the probability λ of experiencing "good" 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.

The rationale for increasing inefficiency in the optimal contract follows from the persistent informational linkage between the buyer's private information at the time of contracting and her values in future periods. As first shown by Baron and Besanko (1984), distortions are most effective at reducing information rents at histories where the buyer's value is most affected by her initial type. For instance, when values are i.i.d. in each period, the initial type is uninformative about future values and only the initial period is distorted; on the other hand, when values do not change over time, the initial type is perfectly informative and distortions are constant. Since values in our environment are the product of conditionally independent shocks that depend on the buyer's initial type, the impact of λ accumulates with each additional shock; therefore, the distribution of values becomes more sensitive to λ over time. Therefore, distortions increase over time, manifesting in progressive screening and increasingly aggressive exclusion of buyers.

4.3. *Convex costs*

The results presented above extend beyond the unit-demand setting above; a similar contractual structure arises when 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 in each period at a cost of $c(q)=q^2/2$. Then the seller's relaxed problem ($\mathcal{R}'$) becomes

$$\max_{\{\mathbf{q},\mathbf{p}\}}\left\{\sum_{t=1}^{T}\delta^t \iint_{\Lambda\times A^t}\left(\varphi(\alpha^t,\lambda)q_t(\alpha^t,\lambda)-\frac{q_t^2(\alpha^t,\lambda)}{2}\right)\mathrm{d}W^t(\alpha^t|\lambda)\mathrm{d}F(\lambda)\right\}$$

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

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 inferred value resulting from a report of u), thereby decreasing the (efficient) 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 $t=1,\dots,T$ and $s\geq t$,

$$q_s^*(\alpha_{-t}^s,u,\alpha^{t-1},\lambda)\geq q_s^*(\alpha_{-t}^s,d,\alpha^{t-1},\lambda) \text{ for all } \alpha^{t-1}\in A^{t-1},\alpha_{-t}^s\in A^{s-t}, \text{ and } \lambda\in\Lambda.$$

Since this 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 ($t \geq 1$) 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 4.2, 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

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

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 4.2:

$$\begin{aligned} p_0^*(\lambda) := & \sum_{t=1}^{T} \delta^t \int_{A^t} \left(q_t^*(\alpha^t, \lambda) v(\alpha^t) - p_t^*(\alpha^t, \lambda) \right) \mathrm{d}W^t(\alpha^t | \lambda) \\ & - \sum_{t=1}^{T} \delta^t \int_0^{\lambda} \int_{A^t} \bar{q}_t^*(d, \alpha^{t-1}, \mu) \Delta \,\mathrm{d}W^t(\alpha^t | \mu) \mathrm{d}\mu. \end{aligned} \tag{15}$$

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

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

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.[12] Thus, for any initial-period report λ, the contract described in Equations (13)–(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 in Lemma 4.2 is only a necessary implication of period-zero incentive compatibility. However, Condition A implies that the quantity schedules are 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 4.2. *Suppose that the distribution F satisfies Condition 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* $(\mathcal{R}')$.

In this setting, the sufficient condition on F guarantees that q_t^* is monotone increasing in λ; therefore, the seller's menu is infinite. However, as in the indivisible goods case, the optimal contract permits only a fixed finite number of reported d shocks before permanently excluding the buyer, where this upper bound depends only on the buyer's report of 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 $\lambda < 1$, inefficient exclusion is unavoidable given a sufficiently long time horizon T.

5. CONTINUOUS SHOCKS

We now present a more general formulation of the model where, instead of discrete shocks, the buyer's valuation shocks in each period are drawn from a continuous distribution. More specifically, we now assume that $\Lambda := [\underline{\lambda}, \bar{\lambda}]$ with $0 \leq \underline{\lambda} < \bar{\lambda} \leq \infty$. We assume that f, the density of F, is strictly positive and differentiable on Λ. Moreover, we assume that, for all $\lambda \in \Lambda$, the support of the conditional distribution $G(\cdot|\lambda)$ is the interval $A := [\underline{\alpha}, \bar{\alpha}]$ with $0 \leq \underline{\alpha} < \bar{\alpha} \leq \infty$. We assume that G is twice continuously differentiable, and denote by $g(\cdot|\lambda)$ the conditional density of $G(\cdot|\lambda)$. Finally, we assume that $g(\cdot|\lambda)$ is strictly positive on A for all $\lambda \in \Lambda$. Note that we maintain our earlier assumption of first-order stochastic dominance, implying that $\partial G(\alpha|\lambda)/\partial\lambda \leq 0$.

12. 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.

5.1. *Simplifying the seller's relaxed problem*

As in Section 4, our analysis of the seller's relaxed problem ($\mathcal{R}$) begins with simplifying the single-deviation and participation constraints. With continuous shocks, we can use the Mirrlees (1971) first-order approach (with details in the Appendix) to "localize" the period-t constraints:

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 $(\alpha^{t-1},\lambda)\in A^{t-1}\times\Lambda$,*

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

$$\bar{q}_t(\alpha_t,\alpha^{t-1},\lambda)\text{ is nondecreasing in }\alpha_t;\text{ and} \tag{MON$''$-t}$$

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

Recall that the buyer's initial-period private information λ does not directly affect her payoffs. Therefore, the standard single-crossing condition does not apply, and we resort instead to an envelope argument (with proof in the Appendix) to simplify the seller's relaxed problem and remove the payment rules from the objective function. As with discrete shocks, 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 individual rationality constraint (IR-0) *is equivalent to the requirement that*

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

With these results in hand, we return to the seller's problem. Since (IC$''$-0) must hold in any incentive compatible mechanism, standard techniques imply that the relaxed problem ($\mathcal{R}$) becomes

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

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

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 we may write

$$\begin{aligned}\sum_{t=1}^{T}\delta^t\int_{A^t}\bar{q}_t(\alpha^t,\lambda)\frac{\partial G(\alpha_t|\lambda)/\partial\lambda}{g(\alpha_t|\lambda)}\mathrm{d}W^t(\alpha^t|\lambda)\\ =\sum_{t=1}^{T}\delta^t\int_{A^t}q_t(\alpha^t,\lambda)\sum_{s=1}^{t}v(\alpha^t_{-s},\alpha^{s-1})\frac{\partial G(\alpha_s|\lambda)/\partial\lambda}{g(\alpha_s|\lambda)}\mathrm{d}W^t(\alpha^t|\lambda).\end{aligned}\tag{16}$$

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

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

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

5.2. *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 $t=1,\ldots,T$ is

$$\begin{aligned}\varphi(\alpha^t,\lambda):=&v(\alpha^t)+\sum_{s=1}^{t}v(\alpha^t_{-s},\alpha^{s-1})\frac{\partial G(\alpha_s|\lambda)/\partial\lambda}{g(\alpha_s|\lambda)}\frac{1-F(\lambda)}{f(\lambda)}\\ =&v(\alpha^t)+v(\alpha^t)\sum_{s=1}^{t}\frac{1}{\alpha_s}\frac{\partial G(\alpha_s|\lambda)/\partial\lambda}{g(\alpha_s|\lambda)}\frac{1-F(\lambda)}{f(\lambda)}.\end{aligned}\tag{17}$$

The first term in each of these expressions is the buyer's contribution to the social surplus, whereas the second term represents the information rents that must be left to the buyer in order to induce truthful revelation of her private information.[13] 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, the additional $\sum_{s=1}^{t}(v(\alpha^t)/\alpha_s)((\partial G(\alpha|\lambda)/\partial\lambda)/g(\alpha|\lambda))$ term is the continuous shock analog of the summation in Equation (7): it is an "informativeness measure" as in Baron and Besanko (1984) which reflects the persistent informational linkage between λ and future values, where we sum over all $s\le t$ to account for the different shocks through which this influence manifests.

We will now specialize the problem to the case where the seller faces the increasing and convex cost function $c(q)=q^2/2$. Pointwise maximization (for each (α^t,λ) tuple) of the integrand in ($\mathcal{R}''$) while ignoring (for now) the remaining constraints yields the following solution:

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

13. Recall that $\partial G(\alpha|\lambda)/\partial\lambda\le 0$ (due to first-order stochastic dominance), and so these information rents are not paid *by* the buyer, but rather *to* her.

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 (17) yields $\varphi(\alpha^t,\lambda)\neq\varphi(\hat{\alpha}^t,\lambda)$.[14] 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 justify this approach, we require an additional separability assumption on the conditional distribution of shocks that is sufficient for the path-independence of the allocation rule:

Condition B.1. *There exist constants $a,b\in\mathbb{R}$ and a function $\gamma:\Lambda\to\mathbb{R}$ such that, for all $\alpha\in A$ and all $\lambda\in\Lambda$,*

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

Notice that when this condition is satisfied, we may write the buyer's virtual value as

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

that is, 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. Thus, Condition B.1 is a key part of our characterization of environments in which incentives decouple over time.

Clearly, this condition is *not* without loss of generality. However, there are many natural and commonly used parametric classes of distributions that satisfy Condition B.1. For example, when $\alpha=\lambda z$ (where z is an independent random variable drawn from an arbitrary distribution that admits a density), then $(\partial G(\alpha|\lambda)/\partial\lambda)/g(\alpha|\lambda)=-\alpha/\lambda$. Similarly, if the shocks are distributed according to 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 the ratio in question again equals $-\alpha/\lambda$. If G is a lognormal distribution with mean λ and arbitrary nonzero variance, then $(\partial G(\alpha|\lambda)/\partial\lambda)/g(\alpha|\lambda)=-\alpha$. Another example is the power distribution $G(\alpha|\lambda)=(\kappa\alpha)^\lambda$ for $\alpha\in[0,1/\kappa]$ and $\kappa>0$; in this case, $(\partial G(\alpha|\lambda)/\partial\lambda)/g(\alpha|\lambda)=\alpha\log(\kappa\alpha)/\lambda$. Thus, while the class of environments we characterize is restricted, it certainly includes many cases of interest.

Whenever Condition B.1 is satisfied, 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:

$$\hat{q}_t^*(v(\alpha^t),\lambda):=q_t^*(\alpha^t,\lambda)=\max\{\varphi(\alpha^t,\lambda),0\}. \tag{18}$$

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 nonlinear pricing problem: we define

$$p_t^*(\alpha^t,\lambda):=\hat{q}_t^*(v(\alpha^t),\lambda)v(\alpha^t)-\int_{\underline{\alpha}^t}^{v(\alpha^t)}\hat{q}_t^*(v',\lambda)\mathrm{d}v', \tag{19}$$

14. 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 any permutation σ.

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 Condition B.1 implies that incentives for truthful reporting in the period-t mechanism are completely decoupled from the incentives in any other period—the initial-period report of λ determines the 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 (18) to satisfy these single-deviation constraints:

Condition B.2. *For all $t=1,\dots,T$, the allocation rule q_t^* in Equation* (18) *is increasing in $v(\alpha^t)$.*

Moreover, note that—since $\hat{q}_t^*(v(\alpha^t),\lambda)\geq 0$ for all $(\alpha^t,\lambda)\in A^t\times\Lambda$—Equation (19) implies that the buyer's flow utility in each period (when reporting truthfully) is nonnegative. 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 (1) implies that

$$\begin{aligned}
p_0^*(\lambda) &:= \sum_{t=1}^{T}\delta^t \int_{A^t}(q_t^*(\alpha^t,\lambda)v(\alpha^t)-p_t^*(\alpha^t,\lambda))\mathrm{d}W^t(\alpha^t|\lambda)-U_0(\lambda)\\
&= \sum_{t=1}^{T}\delta^t \int_{A^t}\int_{\underline{\alpha}^t}^{v(\alpha^t)}\hat{q}_t^*(v',\lambda)\mathrm{d}v'\mathrm{d}W^t(\alpha^t|\lambda)\\
&\quad +\sum_{t=1}^{T}\delta^t \int_{\underline{\lambda}}^{\lambda}\int_{A^t}\bar{q}_t^*(\alpha^t,\mu)\frac{\partial G(\alpha_t|\mu)/\partial\lambda}{g(\alpha_t|\mu)}\mathrm{d}W^t(\alpha^t|\mu)\mathrm{d}\mu.
\end{aligned} \tag{20}$$

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.

Condition B.3. *For all $t=1,\dots,T$, the allocation rule q_t^* in Equation* (18) *is increasing in λ.*

This condition is the counterpart to Condition A, and the following theorem (which we prove in the Appendix) is the counterpart in this more general setting to Theorems 4.1 and 4.2.

Theorem 5.3. *Suppose that Conditions 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* (18) *and* $\mathbf{p}^*$ *denotes the payment rules from Equations* (19) *and* (20)*, is an optimal contract that solves the seller's problem* ($\mathcal{R}''$).

Having established this result, let us explore the dynamic properties of the optimal contract in our setting. To this end, 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 $(\alpha^{t-1},\lambda)\in A^{t-1}\times\Lambda$, leads to a nonnegative allocation in period t; that is, let

$$k_t(\alpha^{t-1},\lambda):=\inf\left\{\alpha_t'\in A: q_t^*(\alpha_t',\alpha^{t-1},\lambda)>0\right\} \tag{21}$$

for all $t=1,\dots,T$, where we let $k_t(\alpha^{t-1},\lambda):=\bar{\alpha}$ if the set above is empty. Notice that Conditions B.2 and B.3 imply that q_t^* is increasing in λ and α_s for all $s\le t$, so k_t is decreasing in each of its arguments. Therefore, the optimal contract is more permissive for those "lucky" buyers who have experienced relatively high shocks or who have a high value of λ.

Despite this permissiveness, however, the optimal contract in our environment is unforgiving: once the buyer is excluded in period t, she is excluded in all future periods. This is most easily seen using a recursive formulation of the buyer's virtual value: note that the buyer's period-$(t+1)$ virtual value, given $\alpha^t\in A^t$ and $\lambda\in\Lambda$, may be written as

$$\begin{aligned}\varphi(\alpha_{t+1},\alpha^t,\lambda)&=v(\alpha^{t+1})\left(1+\sum_{s=1}^{t+1}\frac{1}{\alpha_s}\frac{\partial G(\alpha_s|\lambda)/\partial\lambda}{g(\alpha_s|\lambda)}\frac{1-F(\lambda)}{f(\lambda)}\right)\\&=\alpha_{t+1}v(\alpha^t)\left(1+\sum_{s=1}^{t}\frac{1}{\alpha_s}\frac{\partial G(\alpha_s|\lambda)/\partial\lambda}{g(\alpha_s|\lambda)}\frac{1-F(\lambda)}{f(\lambda)}\right)+v(\alpha^t)\frac{\partial G(\alpha_{t+1}|\lambda)/\partial\lambda}{g(\alpha_{t+1}|\lambda)}\frac{1-F(\lambda)}{f(\lambda)}\\&=\alpha_{t+1}\varphi(\alpha^t,\lambda)+v(\alpha^t)\frac{\partial G(\alpha_{t+1}|\lambda)/\partial\lambda}{g(\alpha_{t+1}|\lambda)}\frac{1-F(\lambda)}{f(\lambda)}.\end{aligned}\tag{22}$$

Since $\partial G(\alpha|\lambda)/\partial\lambda\le 0$ for all α and λ via first-order stochastic dominance, Equation (22) implies that $\varphi(\alpha_{t+1},\alpha^t,\lambda)\le 0$ for all $\alpha_{t+1}\in A$ whenever $\varphi(\alpha^t,\lambda)\le 0$; equivalently, $k_{t+1}(\alpha^t,\lambda)=\bar{\alpha}$ whenever $\alpha_t\le k_t(\alpha^{t-1},\lambda)$. Therefore, 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, the optimal contract involves a form of "tightening the screws," as the set of reports that lead to a positive quantity in any period $t+1\le T$ is contained in the corresponding set for period t. To see this, suppose that $\alpha_t=k_t(\alpha^{t-1},\lambda)>\underline{\alpha}$, so that the buyer is "just barely" excluded in period t. Since this implies that $\varphi(\alpha_t,\alpha^{t-1},\lambda)=0$, Equation (22) can be rewritten, for any period-$(t+1)$ shock $\alpha_{t+1}>\alpha_t$, as

$$\varphi(\alpha_{t+1},\alpha^t,\lambda)=v(\alpha^t)\frac{\partial G(\alpha_{t+1}|\lambda)/\partial\lambda}{g(\alpha_{t+1}|\lambda)}\frac{1-F(\lambda)}{f(\lambda)}\le 0.$$

Therefore, $q_{t+1}^*(\alpha_{t+1},\alpha^t,\lambda)=0$ and $k_{t+1}(k_t(\alpha^{t-1},\lambda),\alpha^{t-1},\lambda)=\bar{\alpha}$. Thus, a buyer who is on the cusp of allocation in period t is always excluded in period $t+1$.

Finally, recall that $k_{t+1}(\alpha^t,\lambda)$ is decreasing in α_t. This property, combined with the observation above, implies that, for any $\alpha_t\ge k_t(\alpha^{t-1},\lambda)$, 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 $[k_t(\alpha^{t-1},\lambda),\bar{\alpha}]$.

These features of the optimal contract are the continuous analogs of the finite honeymoon phases that arise with discrete shocks and single-unit demand. 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 is shrinking over time. Thus, the seller progressively screens the buyer by restricting supply and increasing the probability of permanent exclusion as the relationship progresses.

6. CONCLUDING REMARKS

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 introduce sufficient conditions on the underlying primitives that allow us to solve for the seller's optimal contract, taking into account the buyer's incentives for participation and for truthful revelation throughout the interaction. These conditions characterize a class of environments in which incentives decouple over time. When this is the case, the optimal long-term contract features surprisingly simple menus of options that vary not only by upfront cost and future strike prices, 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 in exchange for lower strike prices. Moreover, we identify an additional mechanism by which the seller 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 upward, 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 critical assumption in our model is that the buyer's value in each period is the product of a sequence of conditionally independent shocks. This assumption imposes a great deal of structure on the environment; in particular, it implies that shocks have a symmetric impact on values (that is, the buyer's value is a commutative function of shocks) and that the distribution of shocks in any given period does not depend on previous values. These two properties are crucial for the decoupling of incentives over time. When shocks are drawn from different distributions over time or depend directly on previous values, then the solution to the relaxed problem need not be path-independent. Of course, this necessitates a different approach to incentive compatibility than that taken in the present work; Pavan, Segal and Toikka (2011) provide several interesting and useful results in this direction.

The present work sets the stage for several avenues of further inquiry. Recall, for instance, that the optimal contract in our setting is not renegotiation-proof, so our assumption of full commitment power on the part of the seller has substantial bite. Understanding the precise role of commitment is therefore a natural topic for additional investigation. In addition, 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. 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

Lemma A. *For all $\lambda, \lambda' \in \Lambda$, we may write*

$$\widehat{U}_0(\lambda',\lambda) = -p_0(\lambda') + \delta \int_A U_1(\alpha_1,\lambda')\mathrm{d}G(\alpha_1|\lambda) + \sum_{t=2}^{T} \delta^t \int_{A^{t-1}} \int_A U_t(\alpha^t,\lambda')\mathrm{d}[G(\alpha_t|\lambda) - G(\alpha_t|\lambda')]\mathrm{d}W^{t-1}(\alpha^{t-1}|\lambda).$$

Proof Notice that for any $s=1,\ldots,T$, we may write

$$\begin{aligned}
&\sum_{t=s}^{T}\delta^t\int_{A^t}\big(q_t(\alpha^t,\lambda')v(\alpha^t)-p_t(\alpha^t,\lambda')\big)\,\mathrm{d}W^t(\alpha^t|\lambda)\\
&\qquad=\sum_{t=s}^{T}\delta^t\int_{A^t}\big(q_t(\alpha^t,\lambda')v(\alpha^t)-p_t(\alpha^t,\lambda')\big)\,\mathrm{d}W^t(\alpha^t|\lambda)\\
&\qquad\quad+\sum_{t=s+1}^{T}\delta^t\int_{A^t}\big(q_t(\alpha^t,\lambda')v(\alpha^t)-p_t(\alpha^t,\lambda')\big)\,\mathrm{d}W^{t-s}(\alpha^t_{-s}|\lambda')\mathrm{d}W^s(\alpha^s|\lambda)\\
&\qquad\quad-\sum_{t=s+1}^{T}\delta^t\int_{A^t}\big(q_t(\alpha^t,\lambda')v(\alpha^t)-p_t(\alpha^t,\lambda')\big)\,\mathrm{d}W^{t-s}(\alpha^t_{-s}|\lambda')\mathrm{d}W^s(\alpha^s|\lambda)\\
&\qquad=\delta^s\int_{A^s}U_s(\alpha^s,\lambda')\mathrm{d}W^s(\alpha^s|\lambda)-\delta^{s+1}\int_{A^{s+1}}U_{s+1}(\alpha_{s+1},\alpha^s,\lambda')\mathrm{d}G(\alpha_{s+1}|\lambda')\mathrm{d}W^s(\alpha^s|\lambda)\\
&\qquad\quad+\sum_{t=s+1}^{T}\delta^t\int_{A^t}\big(q_t(\alpha^t,\lambda')v(\alpha^t)-p_t(\alpha^t,\lambda')\big)\,\mathrm{d}W^t(\alpha^t|\lambda).
\end{aligned}$$

Therefore, we may rewrite $\widehat{U}_0(\lambda',\lambda)$ from Equation (2) as

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

Substituting in from the expressions above yields

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

Proceeding inductively in this manner, we may conclude that

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

Proof of Lemma 4.1. For any $\alpha_t,\alpha_t'\in A$, $\alpha^{t-1}\in A^{t-1}$ and $\lambda\in\Lambda$, adding and subtracting

$$q_t(\alpha_t',\alpha^{t-1},\lambda)v(\alpha_t',\alpha^{t-1})+\sum_{s=t+1}^{T}\delta^{s-t}\int_{A^{s-t}}q_s(\alpha^s_{-t},\alpha_t',\alpha^{t-1},\lambda)v(\alpha^s_{-t},\alpha_t',\alpha^{t-1})\mathrm{d}W^{s-t}(\alpha^s_{-t}|\lambda)$$

from the right-hand side of the single-deviation constraint (IC-t) yields

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

Letting $\alpha_t = u$ and $\alpha_t' = d$, we can write the constraint 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 4.2. Using Lemma A and the definition of $G(\alpha|\lambda)$, we may write

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

With this in hand, note that

$$\begin{aligned}\frac{\partial}{\partial\lambda}\widehat{U}_0(\lambda',\lambda) = & \sum_{t=1}^{T}\delta^t \int_{A^{t-1}} \Big(U_t(u,\alpha^{t-1},\lambda') - U_t(d,\alpha^{t-1},\lambda')\Big)\,\mathrm{d}W^{t-1}(\alpha^{t-1}|\lambda)\\ & + (\lambda-\lambda')\sum_{t=2}^{T}\delta^t \frac{\partial}{\partial\lambda}\left(\int_{A^{t-1}} \Big(U_t(u,\alpha^{t-1},\lambda') - U_t(d,\alpha^{t-1},\lambda')\Big)\,\mathrm{d}W^{t-1}(\alpha^{t-1}|\lambda)\right).\end{aligned}$$

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

$$U_0'(\lambda) = \frac{\partial}{\partial\lambda}\widehat{U}_0(\lambda',\lambda)\bigg|_{\lambda'=\lambda} = \sum_{t=1}^{T}\delta^t \int_{A^{t-1}} \Big(U_t(u,\alpha^{t-1},\lambda) - U_t(d,\alpha^{t-1},\lambda)\Big)\,\mathrm{d}W^{t-1}(\alpha^{t-1}|\lambda). \qquad \|$$

Proof of Theorem 4.1. Note that we may rewrite $\widehat{U}_0(\lambda',\lambda)$ as

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

where we use the expectations operator $\mathbb{E}[\cdot]$ to economize on notation. Therefore,

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

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

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

where the equality 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

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

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

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

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

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

so we must have

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

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

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

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

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

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

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

Finally, note that

$$\begin{aligned}
&\frac{t}{t-m_t+k+1} \binom{t-1}{m_t-1} \binom{m_t-1}{k} (-1)^k \mu^{t-m_t+k+1} \\
&\quad = \frac{t-m_t+1}{t-m_t+k+1} \binom{t}{m_t-1} \binom{m_t-1}{k} (-1)^k \mu^{t-m_t+k+1} \\
&\quad = \binom{t}{m_t-1-k} \binom{t-m_t+k}{k} (-1)^k \mu^{t-m_t+k+1}.
\end{aligned}$$

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

$$\begin{aligned}
&\sum_{k=0}^{m_t-1} \frac{t}{t-m_t+k+1}\binom{t-1}{m_t-1}\binom{m_t-1}{k}(-1)^k \mu^{t-m_t+k+1} \\
&\qquad = \sum_{k=0}^{m_t-1}\sum_{j=0}^{k}\binom{t}{j+m_t-k-1}\binom{j+m_t-k-1}{j}(-1)^j \mu^{t-m_t+k+1} \\
&\qquad = \sum_{j=0}^{m_t-1}\sum_{k=0}^{j}\binom{t}{j}\binom{j}{k}(-1)^k \mu^{t-j+k} \\
&\qquad = \sum_{j=0}^{m_t-1}\binom{t}{j}\mu^{t-j}(1-\mu)^j.
\end{aligned}$$

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

$$\begin{aligned}
&\int_{\lambda'}^{\lambda}\mathbb{E}\left[\bar{q}_t^*(d,\alpha^{t-1},\lambda')\Delta\middle|\mu\right]\mathrm{d}\mu \\
&\qquad + \mathbb{E}\left[q_t^*(\alpha^t,\lambda')(v(\alpha^t)-p_t^*(\alpha^t,\lambda'))\middle|\lambda'\right] - \mathbb{E}\left[q_t^*(\alpha^t,\lambda')(v(\alpha^t)-p_t^*(\alpha^t,\lambda'))\middle|\lambda\right] = 0.
\end{aligned}$$

Therefore, for all $\lambda,\lambda'\in\Lambda$, $\widehat{U}_0(\lambda,\lambda)\geq\widehat{U}_0(\lambda',\lambda)$; that is, for each $\lambda\in\Lambda$, $\widehat{U}_0(\lambda',\lambda)$ 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 4.2. Note that we may write $\widehat{U}_0(\lambda',\lambda)$ as

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

where the equality comes from the identity in Equation (6) and we use the expectations operator $\mathbb{E}[\cdot]$ to economize on notation. 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

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

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

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

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

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

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

$$\begin{aligned}&\sum_{j=1}^{t}\sum_{k=0}^{j-1}\binom{t}{k}\Big((t-k)(\lambda')^{t-k-1}(1-\lambda')^k-k(\lambda')^{t-k}(1-\lambda')^{k-1}\Big)q_t^*(j,\lambda')\Delta u^{t-j}d^{j-1}\\ &=\sum_{j=0}^{t-1}q_t^*(j+1,\lambda')\Delta u^{t-j-1}d^j\sum_{k=0}^{j}\binom{t}{k}\Big((t-k)(\lambda')^{t-k-1}(1-\lambda')^k-k(\lambda')^{t-k}(1-\lambda')^{k-1}\Big).\end{aligned}$$

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

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

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 $\partial\widehat{U}_0(\lambda',\lambda)/\partial\lambda'$ sum to zero; that is,

$$\begin{aligned}\frac{\partial \widehat{U}_0(\lambda',\lambda)}{\partial\lambda'} &= \sum_{t=1}^{T}\delta^t\sum_{k=0}^{t}\binom{t}{k}\Big(\lambda^{t-k}(1-\lambda)^k-(\lambda')^{t-k}(1-\lambda')^k\Big)\sum_{j=k+1}^{t}\frac{\partial q_t^*(j,\lambda')}{\partial\lambda'}\Delta u^{t-j}d^{j-1}.\\ &=\sum_{t=1}^{T}\delta^t\big(\mathbb{E}[\Phi_t(\kappa_t)|\lambda]-\mathbb{E}[\Phi_t(\kappa_t)|\lambda']\big),\end{aligned}$$

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 $\lambda \in \Lambda$,

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

Thus, holding λ fixed, $\widehat{U}_0(\lambda', \lambda)$ 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 $t \geq 1$, and notice that (IC-t) may be rewritten as

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

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 nonnegative, the integrand above is also nonnegative; 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. Using Lemma A and the fact that $\mathrm{d}G(\alpha|\mu) = g(\alpha|\mu)\mathrm{d}\alpha$ for all $\mu \in \Lambda$ in the continuous–shock setting, we may write

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

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

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

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 nonnegative for all t, implying that $\bar{q}_t$ is also nonnegative 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 5.3. Making use of the definition of $\mathbf{p}^*$ from Equations (20) and (19), we may rewrite $\widehat{U}_0(\lambda',\lambda)$ from Equation (2) as

$$\widehat{U}_0(\lambda',\lambda)=\sum_{t=1}^{T}\delta^t\int_{A^t}\int_{\underline{\alpha}^t}^{v(\alpha^t)}\hat{q}_t^*(v',\lambda')\mathrm{d}v'\,\mathrm{d}W^t(\alpha^t|\lambda)$$

$$-\sum_{t=1}^{T}\delta^t\int_{A^t}\int_{\underline{\alpha}^t}^{v(\alpha^t)}\hat{q}_t^*(v',\lambda')\mathrm{d}v'\,\mathrm{d}W^t(\alpha^t|\lambda')$$

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

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

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

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

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

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

Recall from Equation (16), 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

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

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

Thus, we may conclude that

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

Note, however, that Condition 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 $\lambda \in \Lambda$,

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

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

Finally, note that (as discussed earlier) Conditions 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 $(\mathcal{R}'')$, but also is fully incentive compatible: the buyer prefers truthful reporting to any potential deviation, regardless how complex. ‖

Acknowledgments. We thank Atila Abdulkadiroglu, Dirk Bergemann, Simon Board, Alessandro Bonatti, Jernej Copic, Rahul Deb, Tracy Lewis, Qingmin Liu, Pino Lopomo, Leslie Marx, Hamid Nazerzadeh, Mallesh Pai, Larry Samuelson, Curt Taylor, Juuso Toikka, Joel Watson, and Huseyin Yildirim, as well as conference and seminar participants, for their many helpful comments and suggestions. This work has benefited greatly from the detailed comments of two anonymous referees and the editor, Bruno Biais. 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.

REFERENCES

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

ATHEY, S. and SEGAL, I. (2007b), "An Efficient Dynamic Mechanism", Unpublished manuscript, Harvard University.

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

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

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

BERGEMANN, D. and VÄLIMÄKI, J. (2002), "Information Acquisition and Efficient Mechanism Design", *Econometrica*, **70**, 1007–1033.

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

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

BOARD, S. (2007), "Selling Options", *Journal of Economic Theory*, **136**, 324–340.

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

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

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

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

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

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

GARRETT, D. (2011), "Durable Goods Sales with Dynamic Arrivals and Changing Values", Unpublished manuscript, Northwestern University.

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

GERSHKOV, A. and MOLDOVANU, B. (2009b), "Learning About the Future and Dynamic Efficiency", *American Economic Review*, **99**, 1576–1587.

GERSHKOV, A. and MOLDOVANU, B. (2010a), "Demand Uncertainty and Dynamic Allocation with Strategic Agents", Unpublished manuscript, University of Bonn.

GERSHKOV, A. and MOLDOVANU, B. (2010b), "Efficient Sequential Assignment with Incomplete Information", *Games and Economic Behavior*, **68**, 144–154.

GERSHKOV, A. and SZENTES, B. (2009), "Optimal Voting Schemes with Costly Information Acquisition", *Journal of Economic Theory*, **144**, 36–68.

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

KAKADE, S.M., LOBEL, I. and NAZERZADEH, H. (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**, 719–755.

KRÄHMER, D. and STRAUSZ, R. (2011), "Optimal Procurement Contracts with Pre-Project Planning", *Review of Economic Studies*, **78**, 1015–1041.

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

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

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

MIERENDORFF, K. (2011), "Optimal Dynamic Mechanism Design with Deadlines", Unpublished manuscript, University of Zurich.

MILGROM, P. and SEGAL, I. (2002), "Envelope Theorems for Arbitrary Choice Sets", *Econometrica*, **70**, 583–601.

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

MIRRLEES, J.A. (1971), "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies*, **38**, 175–208.

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

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

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

PAVAN, A., SEGAL, I. and TOIKKA, J. (2011), "Dynamic Mechanism Design: Incentive Compatibility, Profit Maximization, and Information Disclosure", Unpublished manuscript, Massachusetts Institute of Technology.

RIORDAN, M.H. and SAPPINGTON, D.E. (1987a), "Awarding Monopoly Franchises," *American Economic Review*, **77**, 375–387.

RIORDAN, M.H. and SAPPINGTON, D.E. (1987b), "Information, Incentives, and Organizational Mode", *Quarterly Journal of Economics*, **102**, 243–263.

SAID, M. (2012), "Auctions with Dynamic Populations: Efficiency and Revenue Maximization", *Journal of Economic Theory*, forthcoming.

SHI, X. (2012), "Optimal Auctions with Information Acquisition", *Games and Economic Behavior*, **74**, 666–686.

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

WILSON, R.B. (1993), *Nonlinear Pricing*. (Oxford: Oxford University Press)