Informational Robustness in Intertemporal Pricing

Jonathan Libgober and Xiaosheng Mu

Department of Economics, Harvard University

May 22, 2018

Abstract. Consumers may be unsure of their willingness-to-pay for a product if they are unfamiliar with some of its features or have never made a similar purchase before. How does this possibility influence optimal pricing? To answer this question, we introduce a dynamic pricing model where buyers have the ability to learn about their values for a product over time. A seller commits to a pricing strategy, while buyers arrive exogenously and decide when to make a one-time purchase. The seller does not know how each buyer learns about his value for the product, and seeks to maximize profits against the worst-case information arrival processes. We show that a constant price path delivers the (robustly) optimal profit, which is also the optimal profit in an environment where buyers cannot delay. We discuss the role of price-dependent information for this result, and consider an extension with common values and public information.

Keywords. Intertemporal Pricing, Optimal Stopping, Dynamic Information Structures, Robustness, Mechanism Design.

Contact. jlibgober@gmail.com, indefatigablexs@gmail.com. We are particularly indebted to Drew Fudenberg for guidance and encouragement. We also thank Gabriel Carroll, Rahul Deb, Songzi Du, Ben Golub, Jerry Green, Johannes Hörner, Yuhta Ishii, Scott Kominers, Eric Maskin, Tomasz Strzalecki, Juuso Toikka, and various seminar audiences for comments. Any remaining errors are ours.

1. INTRODUCTION

Suppose a monopolist has invented a new durable product, and is deciding how to set prices over time to maximize profit. Consulting the literature on intertemporal pricing,[1] the monopolist would find that keeping the price fixed (at the single-period profit maximizing price) is an optimal strategy when consumers understand the product perfectly and their willingness-to-pay does not vary over time. But a wrinkle arises if consumers may learn something that influences how much they like the product after pricing decisions have been made, a salient issue since the monopolist's product is completely new. For example, when the Apple Watch, Amazon Echo, and Google Glass were released, most consumers had little prior experience to inform their willingness-to-pay. In such a situation, the monopolist might suspect that purchase decisions will depend on the available information–e.g., journalist reviews about the product–which may in turn depend on pricing. The potential for information arrival presents a challenge to the monopolist's problem.

In isolation, components of this setting have been studied extensively. The literature on informative advertising takes as given that there is some information that would inform consumers of their willingness-to-pay (see Bagwell (2007)). In the intertemporal pricing literature, Stokey (1979) recognized that willingness-to-pay may change over time, and that such changes can influence the optimal pricing strategy. And other papers on intertemporal pricing, such as Biehl (2001) and Deb (2014), have used exogenous learning by consumers to motivate their studies of stochastic changes in buyer values.

Despite this apparent interest, we are not aware of any papers that study dynamic pricing while modeling information arrival explicitly. We suspect one reason for this absence relates to technical difficulties. Buyers' purchase decisions depend on the value of information, something that can become intractable in general dynamic environments. While Deb (2014) and Garrett (2016) restore tractability by considering specific evolutions of buyer values, the stochastic processes they consider violate the martingale condition imposed by Bayesian updating. Their approaches are suitable for studying settings with taste shocks, but they do not fully capture learning. So the question of how to price optimally in the face of information arrival is left unanswered.[2]

We introduce a model of intertemporal pricing that incorporates dynamic information arrival, and we demonstrate the optimality of constant price paths in this model. To do this, we adopt

[1] E.g., Stokey (1979), Bulow (1981), Conlisk, Gerstner and Sobel (1984), among others. These papers show that a seller with commitment does not benefit from choosing lower prices in later periods.

[2] One may think that allowing buyers to learn is simply a matter of making them *more patient*, since information arrival provides incentives to delay purchase. By Landsberger and Meilijson (1985), this logic would imply that for any *fixed* information arrival process, a constant price path should be optimal. However, in Appendix D.5 we show that constant prices are not in general optimal without the robust objective we consider here.

the approach of the active literature on robust mechanism design. A seller commits to a pricing strategy, while buyers observe signals of their values, possibly over time, each according to some information structure (or more precisely, information arrival process). We assume that the seller does not know any part of the information arrival processes, and is concerned with the worst possible information structures given the pricing decisions. One justification for this worst-case analysis is that the seller may want to guarantee a good outcome, no matter what the information structures actually are.[3] For our application, another justification would be that an adversary (e.g. a competitor or antagonistic journalist) may be interested in minimizing the seller's profit. If the firm did not have total control over what information consumers might have access to, our framework would be appropriate. As for the commitment assumption, introducing it circumvents issues related to the Coase conjecture. Without seller commitment, this result implies that the worst-case is approximately achieved when buyers know their values, delivering seller profit equal to the minimum buyer valuation.[4]

Our first result is that a longer time horizon does not increase the amount of profit the seller can ensure from each buyer. One explanation is as follows: In each period, the adversarial nature could release information that minimizes the profit in that period.[5] Doing so would make the seller's problem separable across time, eliminating potential gains from intertemporal price discrimination. This intuition is incomplete, because the worst-case information structures for different periods need not be consistent, in the sense that past information may prevent nature from minimizing profits in the future. This feature makes it difficult to find the exact worst case for an arbitrary price path. Instead, we focus on the class of *partitional* information arrival processes. These processes involve the buyer learning whether his value is above or below a given threshold, with this threshold declining over time. We demonstrate that nature can use a partitional information arrival process to hold the seller to a profit no greater than the single-period benchmark.

While the above argument shows that selling only once (at the single-period optimal price) is an optimal strategy with only a single buyer, this pricing strategy forgoes potential future profit when multiple buyers with i.i.d. values arrive over time. In the classic setting with known values, a constant price path maximizes the profit obtained from each arriving buyer, who either buys immediately upon arrival or not at all. This argument does not extend to our problem, since nature can induce delay by promising to reveal information in the future. Such delay could be costly for the seller, due to discounting. However, we show that as nature attempts to convince the buyer to

[3] A more complete discussion of this justification can be found in the robust mechanism design literature, in particular: Chung and Ely (2007), Frankel (2014), Yamashita (2015), Bergemann, Brooks and Morris (2017), Carroll (2015, 2017).
[4] See Section 3.1 for further discussion of the commitment assumption.
[5] For expositional convenience, we think of "nature" choosing the information arrival process to hurt the seller.

delay her purchase, it must increase the probability of purchase to satisfy the buyer's incentives. With constant prices, the profit loss due to delayed sale is always offset by the increased probability of sale. We thus show that a constant price path ensures the greatest worst-case profit, equal to the profit when buyers can only possibly buy upon arrival.

Together, this analysis delivers a result similar to one that has been shown under known values (see e.g. Stokey (1979)): The seller's optimal strategy is to hold the price fixed at the single-period optimal price, and (in the worst-case) buyers purchase either immediately or never. This holds even though the single-period optimum in our problem is different due to buyer learning. In Section 6.1, we extend our main model to nest the known-value setting. Constant prices remain optimal in that extension, suggesting that our results strictly generalize Stokey (1979).

A crucial assumption in our main model is that buyer information in each period can depend on the entire history of realized prices. In Section 7, we consider several variants of the model, which allow for less interaction between prices and information. With only one period, these alternative setups coincide with the single-period models studied by Roesler and Szentes (2017) and Du (2018). Though their one-period profit guarantee is typically higher than ours, we discuss conditions on the information arrival processes that ensure *their* single-period benchmark is still achievable with arriving buyers. We also show in Section 8 that with patient players as well as *common* values and signals, how information depends on prices has vanishing impact on the optimal profit guarantee per buyer.

We begin by reviewing the literature, and then proceed to present the main model. The one-period benchmark is studied in Section 4, and we show that intertemporal incentives do not help the seller in Section 5. Using this result, we demonstrate that constant price paths are optimal in Section 6. Section 7 discusses our timing assumption, while Section 8 presents the extension to common values and public information. Section 9 concludes.

2. LITERATURE REVIEW

This paper is part of an active literature that studies pricing under robustness concerns, where the designer may be unsure of some parameter of the buyer's problem. Informational robustness is a special case, and one that has been studied in static settings. The most similar to our one-period model are Roesler and Szentes (2017) and Du (2018). Both papers consider a setting like ours, where the buyer's value comes from some commonly known distribution, but where the seller does not know the information structure that informs the buyer of her value.[6] Taken together,

[6]Du (2018) extends the analysis to a many-buyer common value auction environment. He constructs a class of mechanisms that extracts full surplus when the number of buyers grows to infinity. The optimal mechanism for

these papers characterize the seller's maxmin pricing policy and nature's minmax information structure in the *static* zero-sum game between them.[7] The one-period version of our model differs from these papers, since we assume that nature can reveal information depending on the *realized* price the buyer faces (see Section 3.1 for further discussion). Moreover, our paper is primarily concerned with dynamics, which is absent from Roesler and Szentes (2017) and Du (2018).

Other papers have considered the case where the value distribution itself is unknown to the seller. For instance, Carrasco et al. (2017) consider a seller who does not know the distribution of the buyer's value, but who may know some of its moments. If the distribution has two-point support, our one-period model becomes a special case of Carrasco et al. (2017) in which the seller knows the support as well as the expected value. But in general, even in the static setting, assuming a prior distribution constrains the possible posterior distributions nature can induce beyond any set of moment conditions.

In our model, nature being able to condition on realized prices is sufficient to eliminate any gains to randomization, even if the randomization is to be done in the future. This may be reminiscent of Bergemann and Schlag (2011), who show that a deterministic price is maxmin optimal (in one period) when the seller only knows the true value distribution to be in some neighborhood of distributions. However, the reasoning in Bergemann and Schlag (2011) is that a single choice by nature yields worst-case profit for all prices. This is not true in our setting, but we are able to construct an information structure for every pricing strategy that shows randomization does not have benefits.

While most of this literature is static, some papers have studied dynamic pricing where the seller does not know the value distribution. Handel and Misra (2014) allow for multiple purchases, while Caldentey, Liu, Lobel (2016), Liu (2016) and Chen and Farias (2016) consider the case of durable goods. As discussed above, information arrival restricts *how* the value evolves, and rules out the cases considered in the literature. In addition, these papers look at different seller objectives; the first three study regret minimization, whereas the last one looks at a particular mechanism that approximates the optimum.

The literature on robust mechanism design has been able to provide optimality foundations for certain simple mechanisms, which tend to be observed in practice. For instance, Carroll (2017) shows how uncertainty over the correlation between a buyer's demand for different goods

finitely many buyers is solved in the special case of two buyers and two value types by Bergemann, Brooks and Morris (2016), and in the general case by Brooks and Du (2018).

[7] Roesler and Szentes (2017) actually motivate their model as one where the *buyer* chooses her optimal information structure; they show that the solution also minimizes the seller's profit. See Appendix D.4 for a related interpretation of our model.

leads the seller to price the goods independently.[8] In the moral hazard setting considered by Carroll (2015), uncertainty over the mapping from an agent's actions into output favors linear compensation schemes. At the moment, however, this literature has had less to say about dynamic environments. Important exceptions are Penta (2015) and Chassang (2013), but these are both rather different from our setting.[9]

Several intertemporal pricing papers (absent robustness concerns) allow for the value to change over time without explicitly modeling information arrival. Stokey (1979) assumes the value changes deterministically given the initial type. Deb (2014) assumes the value is independently redrawn upon Poisson shocks. For Garrett (2016), the value follows a two-type Markov-switching process. As mentioned above, these papers do not impose the martingale condition for expectations. We are not aware of how to solve the buyer's optimal stopping problem under an arbitrary information arrival process. But the maxmin objective allows us to focus on simple and intuitive information structures, making the buyer's problem tractable.

Finally, the closely related literature on information design has also begun to study dynamics (see Ely, Frankel and Kamenica (2015) and Ely (2017)). While we are ultimately concerned with pricing strategies, our work connects to information design because we describe how receiver (buyer) behavior varies depending on how sender (nature) chooses the information structure. Several of our results—in particular, the proof of Lemma 2—bear resemblance to this literature, and they may be of interest outside of our setting.

3. MODEL

A seller (he) sells a durable good at times $t = 1, 2, \ldots, T$, where $T \leq \infty$. In each period t, a single buyer (she) arrives.[10] We let t denote calendar time, and let a index a buyer's arrival time. All parties discount the future at rate δ. The product is costless for the seller to produce,[11] while each buyer has unit demand. We assume that each buyer has (undiscounted) lifetime value v_a from purchasing the object, where v_a is drawn from a distribution F and fixed over time; when there is no confusion, we will omit the subscript and simply write the value as v. The prior distribution

[8]The general link between dynamic allocations and multi-dimensional screening has been long noted in Bayesian settings (see e.g. Pavan et al. (2014) for discussion). While it is interesting that we obtain a result similar to Carroll (2017), our focus on information arrival and single-object purchase is a significant difference.

[9]Penta (2015) considers the dynamic implementation of social choice functions, and Chassang (2013) shows how dynamics enable a principal to approximate robust contracts that may be infeasible under liability constraints.

[10]Our analysis is unchanged if the number of arriving buyers varies over time, provided the value distribution is fixed.

[11]Introducing a cost of c per unit does not change our results: It is as if the value distribution F were "shifted down" by c, and the buyer might have a negative value. The transformed distribution G in Definition 1 below would also be shifted down by c.

F is common knowledge, with support on $\mathbb{R}_+$ and $0 < \mathbb{E}[v] < \infty$. Until Section 8, we assume different buyers have *independent* values.

However, buyers do not directly know their v; instead, they learn about it through signals that arrive over time, via some information structure. To be precise, a *dynamic information structure* (*or information arrival process*) $\mathcal{I}_a$ for a buyer arriving at time a consists of:

- A set of possible signals for every time $t \geq a$, i.e., a sequence of sets $(S_t)_{t=a}^T$, and
- Probability distributions given by $I_{a,t} : \mathbb{R}_+ \times S_a^{t-1} \times P^t \to \Delta(S_t)$, for all t with $a \leq t \leq T$.

Without loss of generality, we assume that all buyers are endowed with the same signal sets S_t, although each one privately observes any particular signal realization. To avoid measurability issues, each signal set S_t is assumed to be at most countably infinite.

We highlight that in the above definition, the distribution of the signal s_t at time t could depend on the buyer's true value $v_a \in \mathbb{R}_+$, the history of her previous signal realizations $s_a^{t-1} = (s_a, s_{a+1}, \ldots, s_{t-1}) \in S_a^{t-1}$, as well as the history of *all previous and current* prices $p^t = (p_1, p_2, \ldots, p_t) \in P^t$. With independent buyer values, the seller's profit can be minimized on a per buyer basis. Thus there is no need to correlate information across buyers, or to condition a buyer's signal on the purchase history of previous buyers.

The timing of the model is as follows. At time 0, the seller commits to a pricing strategy σ, which is a distribution over possible price paths $p^T = (p_t)_{t=1}^T$. We allow $p_t = \infty$ to mean that the seller refuses to sell in period t. Note that the price the seller posts at time t must be the same for all buyers that have arrived and not yet purchased.[12] After the seller chooses the strategy, nature chooses a dynamic information structure for each buyer. In each period $t \geq 1$, the price in that period p_t is realized according to $\sigma(p_t \mid p^{t-1})$. A buyer arriving at time a with true value v_a observes the signal s_t with probability $I_{a,t}(s_t \mid v_a, s_a^{t-1}, p^t)$ and decides whether or not to purchase the product.

Given the pricing strategy σ and the information structure $\mathcal{I}_a$, the buyer arriving at time a faces an optimal stopping problem. Specifically, she chooses a stopping time τ_a^* adapted to the joint process of prices and signals, so as to maximize the expected discounted value less price:

$$\tau_a^* \in \operatorname*{argmax}_{\tau} \mathbb{E}\left[\delta^{\tau-a}(\mathbb{E}[v_a|s_a^\tau, p^\tau] - p_\tau)\right].$$

The inner expectation $\mathbb{E}[v_a|s_a^\tau, p^\tau]$ represents the buyer's expected value conditional on realized prices and signals up to and including period τ. The outer expectation is taken with respect to the

[12] Otherwise, multiple buyers do not introduce any extra difficulty beyond the case of a single buyer.

evolution of prices and signals. We allow the stopping time τ_a to take any positive integer value $\leq T$, or $\tau_a = \infty$ to mean the buyer never buys.

The seller evaluates payoffs as if the information structures chosen by nature were the worst possible, given his pricing strategy σ and buyers' optimizing behavior. Hence the seller's payoff is:

$$\sup_{\sigma \in \Delta(p^T)} \inf_{(\mathcal{I}_a),(\tau_a^*)} \sum_{a=1}^{T} \mathbb{E}[\delta^{\tau_a^* - a} p_{\tau_a^*}] \text{ s.t. } \tau_a^* \text{ is optimal given } \sigma \text{ and } \mathcal{I}_a \text{ for each } a.$$

Note that when a buyer faces indifference, ties are broken against the seller. Breaking indifference in favor of the seller would not change our results, but would add cumbersome details.[13]

3.1. Discussion of Assumptions

Several of our assumptions are worth commenting on. First, following the robust mechanism design literature, we assume that the buyer has perfect knowledge of the information structure whereas the seller does not. More precisely, each buyer knows her information structure, and is Bayesian about what information will be received in the future. In contrast, the seller is uncertain about the information structure itself. Our interpretation is that the buyer understands what information she will have access to; for instance, she may always use some product review website and hence know very well how to interpret the reviews.[14] The seller, on the other hand, knows that there are many possible ways buyers can learn, and wants to do well against all these possibilities. In Section 6.1, we will show that our results extend even if the seller knows that buyers begin with at least some prior information. Thus, a deterministic constant price path remains optimal when nature is constrained to provide some particular information (but could provide more) in the first period.

Second, we assume that the value distribution is common knowledge. This restriction is for simplicity, allowing us to focus on information arrival and learning. The assumption also enables us to compare our results to the classic literature on intertemporal pricing. In fact, the known-value setting can be seen as an extreme case of our extended model in Section 6.1.

Third, we assume that the seller *commits* to a pricing strategy. The commitment assumption

[13]When ties are broken against the seller, it follows from our analysis that the $\sup \inf$ is achieved as $\max \min$. This would not be true if ties were broken in favor of the seller.

[14]While it may be a strong assumption that buyers perfectly know the signal distribution far into the future, our results do not rely on extra rationality of the buyers beyond what is typically assumed in *static* robust mechanism design. Specifically, our analysis is unchanged if buyers are instead maxmin over future information, so long as they know how to interpret signals in the current period. Developing that extension requires a conceptual framework separate from the current paper, so we omit the details.

avoids certain technical difficulties related to formalizing learning under ambiguity (see Epstein and Schneider (2007)). In practice, firms like Amazon and Apple are widely followed by consumers and industry experts, meaning that they are able to credibly announce and stick to consistent pricing strategies. And while some strategies may be difficult for a seller to commit to, constant price paths are significantly simpler to implement since deviations are straightforward to detect. On the other hand, we *restrict* the seller to using pricing mechanisms, and rule out for instance mechanisms that randomly allocate the object as a function of reports. We view this as a restriction on the environment, but one that is natural in our main applications of interest where prices are typically utilized. This restriction also allows us to avoid difficulties in working with general dynamic mechanisms, where agent types must capture all future information.

Finally, our key timing assumption is that the information structure in each period is determined *after* the price for that period has been realized. As discussed in the literature review, if the information structure is determined *before* the price is realized, then our one-period model would coincide with Roesler and Szentes (2017) and Du (2018). The question of timing is more delicate under dynamics. Although we believe our main model to be the most natural setup, we consider several alternative models in Section 7, which generalize Roesler and Szentes (2017) and Du (2018) to the dynamic setting. In any event, we think that information could depend (at least somewhat) on price in practice. When shopping online, a buyer's information about a product depends on how prominently it is displayed in the search results. If she sorts products by how expensive they are, then the information structure will be price-dependent.

4. SINGLE PERIOD ANALYSIS

We start with the case where the seller does not worry about intertemporal incentives. We do this by taking $T = 1$, although the results would be identical if buyers were constrained to purchase only upon arrival (or never). To solve this problem, we will define a transformed distribution of the prior F. For expositional simplicity, the following definition assumes F is continuous. All of our results in this paper extend to discrete distributions, though the general definition requires additional care and is relegated to Appendix A.

Definition 1. *Given a continuous distribution F, the transformed distribution G is defined as follows. For $y \in \mathbb{R}_+$, let $L(y)$ denote the conditional expectation of $v \sim F$ given $v \leq y$. Then G is the distribution of $L(y)$ when y is drawn according to F. We call G the "pressed" version of F.*

The pressed distribution G is useful because for any realized price p, nature can only ensure that the object remains unsold with probability $G(p)$. To see this, first observe that any information structure is outcome-equivalent to another that directly recommends one of two actions: To purchase the good or not. Given this simplification, the worse-case information structure must have the following property: As long as the buyer is recommended to buy with positive probability, the buyer who is recommended *not* to buy must have expected value exactly p. Otherwise nature could adjust its recommendation to further decrease the probability of sale.

Finally, subject to the constraint that a buyer who does not buy has fixed expected value (in our case, p), one can show that *partitional information structures* maximize the probability of this recommendation (see e.g. Kolotilin (2015)). In a partitional information structure, the buyer is told whether her value is above or below a certain threshold. Using the above definition of G, we argue that the threshold must be $F^{-1}(G(p))$, making $1 - G(p)$ the probability of sale.

These remarks give us the following proposition:

Proposition 1. *In the one-period model, a maxmin optimal pricing strategy is to charge a deterministic price p^* that solves the following maximization problem:*

$$p^* \in \underset{p}{\operatorname{argmax}}\, p(1 - G(p)). \tag{1}$$

We call p^* the one-period maxmin optimal price and similarly $\Pi^* = p^*(1 - G(p^*))$ the one-period maxmin profit.

It is worth comparing the optimization problem (1) to the standard model without informational uncertainty. If the buyer knew her value, the seller would maximize $p(1 - F(p))$. In our setting, the difference is that the pressed distribution G takes the place of F. This analogy will be useful for the analysis in later sections.

The following example illustrates:

Example 1. *Let $v \sim$ Uniform[0,1], so that $G(p) = \min\{2p, 1\}$. Then $p^* = \frac{1}{4}$ and $\Pi^* = \frac{1}{8}$. With only one period to sell the object, the seller charges a deterministic price $1/4$. In response, nature chooses an information structure that tells the buyer whether or not $v > 1/2$.*[15]

In this example, relative to the case where the buyer knows her value, the seller charges a lower price and obtains a lower profit under informational uncertainty. In Appendix D.1, we show that this comparative static holds generally.

[15] There are other information structures that induce the same worst-case profit for the seller; for instance, nature can fully reveal the value when it is above the threshold $1/2$, since the buyer will buy anyways. Nonetheless, the lowest element of the partition cannot be further refined. That is, a buyer whose value is *below* the critical threshold will be told so in every worst-case information structure.

5. INTERTEMPORAL INCENTIVES DO NOT HELP

In this section we present our first main result, that having multiple periods to sell does not allow the seller to extract more surplus from any buyer.[16] Stokey (1979) proved this result for the known-value case, provided buyer value does not change over time. On the other hand, she also demonstrated that if value does change over time, letting the buyer delay purchase could enable the seller to obtain higher profits by facilitating price discrimination. One may wonder whether information arrival, which affects the buyer's expected value over time, could similarly make price discrimination worthwhile. In Appendix D.5, we provide a simple example (with an information arrival process *known* to the seller) where this is the case.

However, it turns out these concerns do not arise for worst-case information structures. Consider the seller's profit from the first buyer. The seller could always sell exclusively in the first period and ensure Π^* as a lower bound. To show that Π^* is also an upper bound, we explicitly construct a dynamic information structure for any pricing strategy, such that the seller's profit under this information structure decomposes into a convex combination of one-period profits. Our proof takes advantage of the partitional form of worst-case information structures from the single-period problem:

Proposition 2. *For any pricing strategy $\sigma \in \Delta(p^T)$, there is a dynamic information structure $\mathcal{I}$ and a corresponding optimal stopping time τ^* that lead to expected (undiscounted) profit no more than Π^* per buyer.*

We focus on the first buyer and show that the seller's worst-case profit from this buyer is at most Π^*. We will present the proof under an additional assumption that the seller charges a *deterministic* price path $(p_t)_{t=1}^T$. This is *not* without loss, because random prices in the future may make it more difficult for nature to choose an information structure in the current period that minimizes profit. However, our argument does extend to random prices and shows that randomization does not help the seller. We discuss this after the more transparent proof for deterministic prices.

Let us first review the sorting argument when the buyer knows her value. In this case, given a price path $(p_t)_{t=1}^T$, we can find time periods $1 \leq t_1 < t_2 < \cdots \leq T$ and value cutoffs $w_{t_1} > w_{t_2} > \cdots \geq 0$, such that the buyer with $v \in [w_{t_j}, w_{t_{j-1}}]$ optimally buys in period t_j (see e.g. Stokey (1979)). This implies that in period t_j, the object is sold with probability $F(w_{t_{j-1}}) - F(w_{t_j})$.

[16] We comment that the dynamics of information arrival are crucial for this result. For instance, suppose the seller knew that information would *not* be released in some period t. Then he could sell exclusively in that period and (by charging random prices) obtain the Roesler and Szentes (2017) profit level, which is generally higher than Π^* (see Section 7 for details). For δ sufficiently close to 1, this pricing strategy does better than a constant price path.

Inspired by the one-period problem, we construct an information structure under which in period t_j, the object is sold with probability $G(w_{t_{j-1}}) - G(w_{t_j})$ (that is, where the pressed distribution G replaces F). The following information structure $\mathcal{I}$ has this property:

- In each period t_j, the buyer is told whether or not her value is in the lowest $G(w_{t_j})$-percentile.
- In all other periods, no information is revealed.

This information structure is similar to the one-period problem, in that a buyer is told whether her value is above or below a threshold. In the dynamic setting, this threshold $F^{-1}(G(w_{t_j}))$ is now declining over time. We refer to any such dynamic information structure as a *partitional information arrival process*, since different signal realizations partition the support of the buyer's value distribution into disjoint intervals. Note that the thresholds are chosen to make the buyer indifferent between purchasing and *continuing without further information.* The buyer therefore prefers to delay purchase when her value is below the threshold. On the other hand, a buyer whose value is above the threshold does not expect to receive further information, and hence purchases immediately. These observations are summarized in the following lemma:

Lemma 1. *Given prices $(p_t)_{t=1}^T$ and the information structure $\mathcal{I}$ constructed above, an optimal stopping time τ^* involves the buyer buying in the first period t_j when she is told her value is **not** in the lowest $G(w_{t_j})$-percentile.*

The formal proof can be found in Appendix A, where we prove a general result for random prices.

Using this lemma, we can now prove Proposition 2 by computing the seller's profit under the information structure $\mathcal{I}$ and the stopping time τ^*:

Proof of Proposition 2 for Deterministic Prices. We assume $T = \infty$, but the same proof works for finite T (with a minor modification to the Abel summation formula used below). Since the buyer with true value v in the percentile range $(G(w_{t_j}), G(w_{t_{j-1}})]$ buys in period t_j, the seller's discounted profit is given by

$$\begin{aligned}\Pi &= \sum_{j\geq 1} \delta^{t_j-1} p_{t_j} \cdot \big(G(w_{t_{j-1}}) - G(w_{t_j})\big) \\ &= \sum_{j\geq 1} (\delta^{t_j-1} p_{t_j} - \delta^{t_{j+1}-1} p_{t_{j+1}}) \cdot (1 - G(w_{t_j})) \\ &= \sum_{j\geq 1} (\delta^{t_j-1} - \delta^{t_{j+1}-1}) w_{t_j} \cdot (1 - G(w_{t_j})) \\ &\leq \delta^{t_1-1} \cdot \Pi^*,\end{aligned} \tag{2}$$

where the second line is by Abel summation,[17] the third line is by type w_{t_j}'s indifference between buying in period t_j or t_{j+1}, and the last inequality uses $w_{t_j}(1-G(w_{t_j})) \le \Pi^*, \forall j$. ■

Relative to the potential complexity of an arbitrary information arrival process, the partitional information structures constructed above are intuitive: Consumers buy when they find out that their value is above some (price-contingent) threshold. Intertemporal pricing cannot help the seller as long as he is concerned at least with this special class of information arrival processes.

Despite the analogy to the known-value case, we highlight that for an arbitrary declining price path, the parititional information structures considered in our proof may not be the worst case (even among partitional processes). The following example illustrates:

Example 2. *Let $T = 2$, $v = 0$ or 1 with equal probabilities, and $\delta = 1/2$. Suppose the seller sets prices to be $p_1 = 11/40$ and $p_2 = 1/10$. Under these prices, a buyer with value $\frac{9}{20}$ would be indifferent (in the first period) between purchase and delay. Hence the partitional information structure constructed in Lemma 1 induces expected value $\frac{9}{20}$ when recommending the buyer not to purchase in the first period. The information structure further induces expected value $p_2 = 1/10$ when recommending the buyer not to purchase in the second period either.*

If the probability of being recommended to buy in period t (conditional on not having bought) is r_t, we have $\frac{1}{2} = r_1 + \frac{9}{20}(1-r_1)$ and $\frac{9}{20} = r_2 + \frac{1}{10}(1-r_2)$. These equations give $r_1 = \frac{1}{11}$ and $r_2 = \frac{7}{18}$. Hence the seller's expected profit against this information structure is

$$p_1 \cdot \frac{1}{11} + (\delta p_2) \cdot \left(1 - \frac{1}{11}\right)\left(\frac{7}{18}\right) \approx 0.0427 < 0.0858 \approx \Pi^*.$$

Now suppose that instead, nature were to provide no information in the first period and reveal the value perfectly in the second period. Note that the buyer would be willing to delay, since

$$\mathbb{E}[v] - p_1 \le \delta \cdot \mathbb{P}[v=1]\,(1-p_2)\,,$$

which in fact holds with equality. Under this alternative information structure, the seller's profit is therefore $\delta p_2 \mathbb{P}[v=1] = \frac{1}{40} < 0.0427$.

The important feature of the example is that by promising more information to the buyer in the second period, nature can create option value and induce delay. This turns out to hurt the seller's expected profit when prices decrease over time, although we show in the next section that

[17] Abel summation says that $\sum_{j\ge1} a_j b_j = \sum_{j\ge1}\left((a_j - a_{j+1})\sum_{i=1}^{j} b_i\right)$ for any two sequences $\{a_j\}_{j=1}^{\infty}, \{b_j\}_{j=1}^{\infty}$ such that $a_j \to 0$ and $\sum_{i=1}^{j} b_i$ is bounded. We take $a_j = \delta^{t_j - 1} p_{t_j}$ and $b_j = G(w_{t_{j-1}}) - G(w_{t_j})$.

the seller can be guarded against such damage with non-decreasing prices.

We mention that in the particular example above, the (latter) information structure we considered is indeed the worst case. However, it is in general challenging to characterize the worst case against a given decreasing price path. As that is not necessary for our main result on the optimality of constant prices, we leave the characterization for future work.

To conclude the section, we briefly discuss how random prices complicate our argument. When prices are random, the threshold values w_{t_j}, if defined using the buyer's indifference condition, will be random variables. A technical difficulty arises because these thresholds may not be monotonically decreasing. When such non-monotonicity occurs, we will not be able to express the seller's discounted profit as a convex sum of one-period profits, and the profit bound in (2) will not be valid.

In Appendix A we show that the basic intuition from the deterministic case extends to random prices, but we need additional tools to generalize the construction appropriately. Specifically, we modify the relevant indifference thresholds so that they are *forced* to be decreasing. Let v_t be the smallest value (in the known-value case) that is indifferent between buying in period t at price p_t and optimally stopping in the future, and *then* let $w_t = \min\{v_1, v_2, \ldots, v_t\}$. We think of this as keeping track of the "binding" thresholds, above which the buyer has already purchased. This circumvents the potential non-monotonicity issue, and we can use the re-defined w_ts to specify the otherwise identical partitional information structure. The rest of the proof proceeds as before, with the assistance of a key lemma (Lemma 4) that expresses the price as the expectation of present and discounted future threshold w_ts. This identity replaces the indifference condition we utilized to derive the third line of (2). Proposition 2 thus continues to hold for random prices.

6. OPTIMALITY OF CONSTANT PRICES

We now demonstrate the optimality of constant price paths. By Proposition 2, the seller's discounted profit from the buyer arriving at time a is bounded above by $\delta^{a-1} \cdot \Pi^*$. This gives us an upper bound for the seller's overall worst-case profit. In the other direction, if the seller were able to set personalized prices, this upper bound could be achieved by selling only once to each arriving buyer. We will show that the seller can achieve the same profit level by always charging p^*, without conditioning prices on the arrival time.

Under known values, any arriving buyer facing a constant price path would buy immediately or never, due to impatience. In contrast, the promise of future information in our setting may induce the buyer to delay, even with constant prices. Nevertheless, in the following lemma, we show that against non-decreasing price paths, nature cannot hurt the seller more than providing information *only* upon arrival. So by committing to never lowering the price, the seller obtains the single-period profit guarantee from each buyer.

Lemma 2. *In the multi-period model with only the first buyer, the seller can guarantee* Π^* *with any deterministic price path* $(p_t)_{t=1}^T$ *satisfying* $p^* = p_1 \leq p_t, \forall t$.

We present the intuition here and leave the formal proof to Appendix A. Fixing a non-decreasing price path and an arbitrary dynamic information structure, we consider an alternative information structure that gives a single recommendation to the buyer (to purchase or not) in the first period. The probability that the buyer is recommended to purchase at time 1 in this replacement information structure leaves the *discounted* probability of sale unchanged. In other words, we "push and discount" nature's recommendation to the buyer's arrival time.

Our proof shows that for non-decreasing prices, the first buyer would follow the recommendations of this replacement information structure, while the seller's profit is weakly decreased. Since the seller receives at least Π^* under any information structure that releases information only in the first period, we obtain the lemma. Note that Example 2 shows this argument relies upon non-decreasing prices.

Armed with this lemma, we can show our main result of the paper. The proof is straightforward given our discussions.

Theorem 1. *The seller can guarantee* $\Pi^* \cdot \frac{1-\delta^T}{1-\delta}$ *with a constant price path charging* p^* *in every period. This deterministic pricing strategy is maxmin optimal, and it is uniquely optimal whenever the one-period maxmin optimal price* p^* *is unique.*

Against a constant price path, a worst-case dynamic information structure simply gives each buyer the same information she would have obtained with only one period to purchase. This completes our analysis of the main model.

6.1. Initial Information

Before proceeding, we point out one extension of our model where constant price paths remain optimal. So far we have assumed that the seller has no knowledge over what information buyers receive. But in practice, the seller may know that buyers observe some specific signals. For example, he may conduct an advertising campaign, and understand its informational impact very well. In that case, the seller would only seek robustness against a subset of the possible information arrival processes.

This situation can be modeled by assuming that in addition to having the prior belief F, each buyer observes some signal $s_0 \in S_0$ before arrival. The seller does not observe the realization of s_0, but the signal set S_0 and the distribution of s_0 given v are common knowledge. This *initial information structure* is denoted by $\mathcal{H}$. We allow nature to provide information conditional on s_0 but keep all other aspects of the model identical.

Let F_{s_0} be the posterior value distribution following signal s_0, and G_{s_0} be its pressed distribution. The same analysis shows that, against F_{s_0}, the (one-period) worst-case information structure involves each buyer being told whether her value is above or below some threshold. Hence,

Proposition 1'. *In the one-period model where the buyer observes initial information structure $\mathcal{H}$, the seller's maxmin optimal price $p^*_{\mathcal{H}}$ is given by:*

$$p^*_{\mathcal{H}} \in \operatorname*{argmax}_p p(1 - \mathbb{E}[G_{s_0}(p)]). \tag{3}$$

The expectation is taken with respect to the distribution of s_0.

The expression (3) is familiar in two extreme cases: if $\mathcal{H}$ is perfectly informative, then F_{s_0} is the point-mass distribution on s_0. This means $G_{s_0}(p)$ is the indicator function for $p \geq s_0$, so that $\mathbb{E}[G_{s_0}(p)] = F(p)$. In the other extreme, $\mathcal{H}$ is completely uninformative and we return to (1).

We can similarly show that a constant price path is optimal for this extended model:

Theorem 1'. *In the multi-period model where each buyer observes initial information structure $\mathcal{H}$ upon arrival, the seller's (maxmin) discounted average profit per buyer is independent of the time horizon T and the discount factor δ. A constant price of $p^*_{\mathcal{H}}$ guarantees this profit.*

The optimality of constant prices in this result can be viewed as a generalization of Stokey (1979), which corresponds to a perfectly informative initial information structure. The proof directly adapts the proof for our main model, so we omit it from the appendix.

7. TIMING

This section analyzes the implications of our assumption regarding the timing of information arrival relative to pricing. This assumption is captured in how we define dynamic information structures, since we allow them to be contingent on all past prices as well as the current price.

When $T = 1$, the alternative model where information cannot depend on price is studied in Roesler and Szentes (2017) and Du (2018), which together solve the optimal selling strategy and the worst-case information structure.[18] For completeness, we recall their result. To make

[18]One may be further interested in cases where information interacts somewhat, but not arbitrarily, with the price. We do not pursue this here.

the connection with our paper most clear, we impose as in these papers that the buyer's value distribution F is supported on $[0, 1]$. Roesler and Szentes (2017) observe that in choosing an information structure, nature is equivalently choosing a distribution $\tilde{F}$ of posterior expected values, such that F is a mean-preserving spread of $\tilde{F}$.[19] They solve for the worst-case distribution $\tilde{F}$ as summarized below:

Theorem 1 in Roesler and Szentes (2017). *For $0 \leq W \leq B \leq 1$, consider the following distribution that exhibits unit elasticity of demand (with a mass point at $x = B$):*

$$F_W^B(x) = \begin{cases} 0 & x \in [0, W) \\ 1 - \frac{W}{x} & x \in [W, B) \\ 1 & x \in [B, 1] \end{cases} \tag{4}$$

In the one-period zero-sum game between the seller and nature, an optimal strategy by nature is to induce posterior expected values given by the distribution F_W^B, such that F is a mean-preserving spread of F_W^B, and W is smallest possible subject to this constraint.

The seller's optimal single-period profit guarantee is equal to the smallest W defined above, which we denote by Π_{RSD}. Conversely, Du (2018) constructs a particular mechanism the seller can use to guarantee profit Π_{RSD} against any information structure. While Du (2018) allows buyers to choose an allocation probability other than 0 or 1, for a single buyer this turns out to generate the same outcome as a *random price mechanism* (see Appendix B.1 for details). We note that $\Pi_{RSD} \geq \Pi^*$ holds, and in Appendix D.3 we characterize when the inequality is strict.

As alluded to in the introduction and Section 3.1, specifying the role of prices in dynamic information structures is more subtle than in a single period. Over time, there are many more ways for information to interact with price. Our main model provides the most cautious profit guarantee, but one may also be interested in how Roesler and Szentes (2017) and Du (2018) extend to dynamic settings.

In this section, we re-define a *dynamic information structure* to be a sequence of signal sets $(S_t)_{t=1}^T$ and probability distributions $I_{a,t} : \mathbb{R}_+ \times S_a^{t-1} \times P^{t-1} \to \Delta(S_t)$. The crucial distinction from our main model is that the signal s_t may depend on previous prices p^{t-1} but not on the current price p_t. Thus the dynamic information structures in this section are a subset of those considered in our main model.

As a warm up, we note that under the alternative setup considered here, a longer horizon does

[19] This equivalence is separately observed by Gentzkow and Kamenica (2016) in the context of Bayesian persuasion. The result appears in the early work of Rothschild and Stiglitz (1970).

not affect the seller's problem with a single buyer.

Proposition 3. *Consider the re-defined class of dynamic information structures. For any T and δ, the seller's maxmin profit from the first buyer (against this class) is given by Π_{RSD}.*

The reasoning is as follows: With multiple periods and a single buyer, the seller can guarantee Π_{RSD} by selling only once in the first period (using Du's mechanism). On the other hand, suppose nature provides the Roesler-Szentes information structure in the first period and no additional information in later periods. Then the seller faces a *fixed* distribution of values given by F_W^B. By Stokey (1979), selling only once is optimal against this distribution, and the seller's optimal profit is $W = \Pi_{RSD}$. This proves the result.

While both Proposition 2 and Proposition 3 show a longer selling horizon does not help the seller, here the argument is more direct due to the duality between Roesler-Szentes and Du. The above proof shows that nature can provide an information structure such that for *all* pricing strategies, the seller's profit is at most Π_{RSD}. In our baseline setting, however, nature must condition the information structure on realized prices (as well as the expected distribution of future prices) in order to hold profit below Π^*.

The same argument shows that even with arriving buyers, the seller cannot guarantee more than Π_{RSD} from each buyer. The question remains as to whether Π_{RSD} per buyer is still achievable when buyers arrive over time (and personalized prices are not allowed). We consider three different cases: In the first case, all dynamic information structures as defined in this section are permitted, and we show that Π_{RSD} per buyer is *not* attainable. In the latter two cases, we show that Π_{RSD} can be guaranteed from each buyer when nature is restricted to certain subsets of information structures. Specifically, the profit bound is tight either if signals do not depend on realized prices, or if each buyer receives a single signal upon arrival.[20]

7.1. Case One: Information can arrive over time and depend upon past prices

First, we allow arbitrary dynamic information structures, so long as information in any period depends only on past realized prices. We show that the one-period profit Π_{RSD} *cannot* be guaranteed from each buyer.[21] For simplicity, the following result assumes two periods, though the same qualitative conclusion holds more generally.

[20]Note that the three cases coincide when $T = 1$.

[21]This result may be expected given the discussion of Case Two and Case Three below, since either of those cases requires a *different* generalization of Du's mechanism to achieve Π_{RSD}.

Claim 1. *Consider the model with two periods and one buyer arriving in each period. Assume that* $\Pi_{RSD} > \Pi^*$ *and that Du's mechanism is uniquely maxmin optimal in the one-period problem.*[22] *Then the seller's total discounted profit guarantee is strictly below* $(1 + \delta)\Pi_{RSD}$ *for any* $\delta \in (0, 1)$.

We prove this claim by constructing a specific information structure that nature provides. When a buyer arrives, nature provides her with the Roesler-Szentes information structure. This yields profit at most Π_{RSD} from the second buyer, and similarly from the first buyer if she expects no additional information in the second period. However, we let nature reveal more information in the second period to induce some (first-period) buyers to delay their purchase. In more detail, in the second period nature reveals the value perfectly to any buyer who *would have* purchased in the first period without this additional information. The key technical step of the proof shows that delay occurs with positive probability and hurts the seller. We comment that we are only able to show total profit is strictly below $(1 + \delta)\Pi_{RSD}$ for this specific information structure. Since it is generally not the worst-case information structure for every pricing strategy, we do not know how to solve for the actual maxmin profit in the current model.

Note that in our construction above, the information structure in the second period depends on the previous signal and the realized price in the first period. The possibility for information to be both *dynamic* and *price-dependent* turns out to be crucial for the result of Claim 1, as we show in the cases below.

7.2. Case Two: Information cannot depend on realized prices

Here we assume that information can arrive over time, but it is independent of realized prices. Formally, we restrict to information structures given by $I_{a,t} : \mathbb{R}_+ \times S_a^{t-1} \to \Delta(S_t)$. It turns out that an optimal strategy for the seller is to utilize a constant price path, where the price is randomly drawn according to Du's mechanism (instead of p^* as in our main model).

Theorem 2. *Suppose that information is independent of realized prices. For any* T *and* δ, *the seller's maxmin average profit per buyer is* Π_{RSD}. *This can be achieved by randomizing over constant price paths.*

This theorem involves new techniques that may be of independent interest. Recall that, in accommodating random pricing strategies in Section 5, we defined "cutoff values" in two steps—first using the buyer's indifference condition, and then keeping track of the lowest realized values.

[22] $\Pi_{RSD} > \Pi^*$ is clearly necessary for the result: We have shown in our main model that $(1+\delta)\Pi^*$ can be guaranteed. On the other hand, we impose an extra assumption that Du's mechanism is strictly optimal. This is for technical reasons that we explain in Appendix B, and it may not be necessary for the conclusion. In any event, we show in Appendix D.2 that Du's mechanism is indeed unique for *generic* value distributions F.

Intuitively, these were the relevant binding thresholds, above which consumers would have already bought. Using these cutoff values, we constructed an information structure under which the seller's profit is decomposed into a convex sum of one-period profits.

Inspired by this technique, the proof of Theorem 2 introduces the dual definition of *cutoff prices* for a given, price-independent, information structure. Assuming that the seller uses a constant price path, these cutoff prices enable us to similarly decompose the seller's profit into one-period profits. We then invoke several properties of Du's mechanism to bound the profit from below. Details are left to Appendix B.

7.3. Case Three: Information only upon arrival

Finally, we consider the case where each buyer is endowed with an information structure I_a : $\mathbb{R}_+ \times P^{t-1} \to \Delta(S_a)$, and she receives the *single* signal upon arrival. In this way, we allow information to respond to realized prices, although for each buyer information does not arrive over time.

Here we immediately see that a constant price path does not in general guarantee Π_{RSD}: If prices were perfectly correlated across periods, nature could provide the worst-case partition to all buyers after the first period. Doing so would hold profit down to Π^*, which is generally lower than Π_{RSD}. This observation suggests that the seller should avoid correlated prices. Indeed, we find that the seller can still guarantee Π_{RSD} per buyer by choosing independent prices across periods:

Theorem 3. *Suppose that each buyer only receives information once upon arrival, before the price in that period realizes. For any T and δ, the seller's maxmin profit average per buyer is Π_{RSD}. This can be achieved with a strategy that sets independent prices across periods.*

Our proof shows how to set independent prices while replicating Du's mechanism for each arriving buyer. This is based on a key lemma (Lemma 5) relating the outcome under a static price distribution to that under a dynamic price distribution.

8. COMMON VALUES AND PUBLIC INFORMATION

This section modifies the model in Section 3 by allowing for common values and publicly observed signals. Notice that making one change without the other would leave the seller's problem unaltered. Here we argue that with both modifications, the seller is able to guarantee higher profits. We do this by studying the stylized case of *pure common values* and *perfectly correlated* *signals.*[23]

Specifically, we assume that all buyers have the same value v, which is drawn from F at $t = 0$. Nature chooses a single information arrival process, $\mathcal{I}$, consisting of signal sets $(S_t)_{1 \le t \le T}$ and distributions $I_t : \mathbb{R}_+ \times S^{t-1} \times P^t \to \Delta(S_t)$. As in our main model, signals can depend on past as well as current prices. However, signals are "public," so a buyer who arrives at time t observes all signals up to period t. Each buyer decides when to purchase based on the signals she observes, and it does not matter whether she also observes others' purchase decisions, which reveal no more information than the signals.[24]

Our first result restricts the set of relevant information arrival processes we need to consider. It turns out that against increasing price paths, it is sufficient for nature to provide a single signal at time 1, which is observed by all buyers:

Lemma 3. *Consider the model with common values and public signals. Fix any price path* $(p_1, \ldots, p_T)$ *with* $p_1 \le p_2 \le \cdots \le p_T$. *Then the worst-case profit is achievable by an information structure that provides a single public signal in the first period.*

This result is the analog of Lemma 2 for the current setting. We use a similar argument to "push and discount" nature's recommendation. The new difficulty here is to ensure that the resulting replacement information structure is still *public*.

Lemma 3 implies that for increasing price paths, we can restrict attention to posterior expected values that are fixed over time. Let $\tilde{F}$ denote the distribution of posterior valuations arising from some information structure. Then the seller's discounted profit can be rewritten as:

$$(1-\delta)\Pi^C = \min_{\tilde{F}} \sum_{t=1}^{T} (1-\delta)\delta^{t-1} p_t \cdot (1 - \tilde{F}(p_t)), \tag{5}$$

Observe that if nature provides the Roesler-Szentes information structure, then the resulting $\tilde{F}$ satisfies $p \cdot (1 - \tilde{F}(p)) \le \Pi_{RSD}$ for every p. Hence the RHS of (5) is at most $(1 - \delta^T)\Pi_{RSD}$.[25] Below we show this upper bound is tight in the infinite-horizon limit with patient players.

[23] Optimal pricing when information is conveyed across buyers has been studied in several other papers in the Bayesian setting, such as Bose et al. (2006, 2008). The distinction is that we allow buyers to delay purchase.

[24] If purchase decisions are observed and values (or signals) are imperfectly correlated, then buyers will face a coordination problem in choosing when to purchase. In general there may be multiple equilibria. However, these strategic considerations among the buyers are not the focus of this paper, and we leave them for future study.

[25] By Stokey (1979), this is an upper bound on seller's profit even if prices are not increasing.

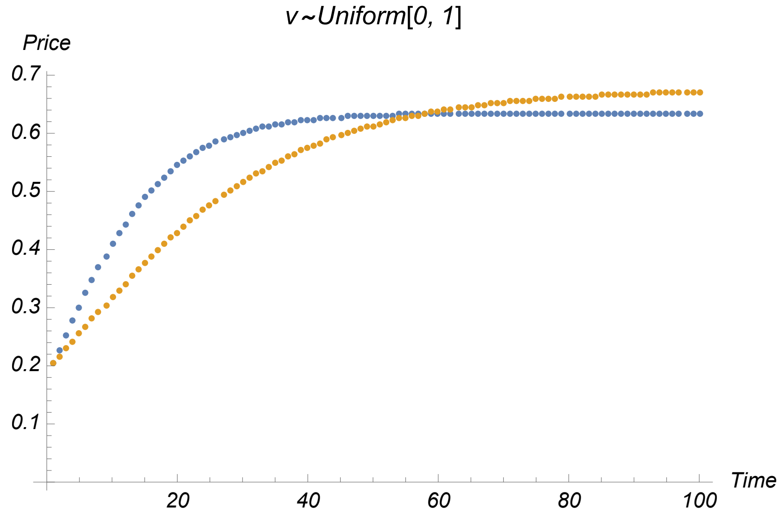

Figure 1: Illustration of constructed price paths. Blue is $\delta = 0.9$; Orange is $\delta = 0.95$.

Theorem 4. *Consider the model with common values and public signals. Let $\Pi^C(\delta, T)$ be the seller's optimal profit with discount factor δ and time horizon T. We have:*

$$\lim_{\delta \to 1, T \to \infty} (1-\delta)\Pi^C(\delta, T) = \Pi_{RSD} \tag{6}$$

The order of limits does not matter, and the maxmin profit guarantee is approximated by a sequence of strictly increasing price paths.

The proof adapts Du's random price mechanism to construct a sequence of price paths such that in the limit, the RHS of (5) converges to the single-period profit under Du's mechanism. These price paths, for uniformly distributed values and $\delta = 0.9$ or $\delta = 0.95$, are shown in Figure 1. Prices rise steeply at first, eventually leveling out.

9. CONCLUSION

In this paper, we have studied optimal monopoly pricing with dynamic information arrival while utilizing a robustness approach. With independent buyer values, the monopolist's optimal profit is what he would obtain with only a single period to sell to each buyer, and a constant price path delivers this optimal profit. The inability to condition on a buyer's arrival time therefore imposes no cost on the seller (in our main model). These conclusions depend on our assumption regarding the timing of information release, and we have illustrated how this is the case.

Our paper contributes to a growing literature which employs the maxmin approach in analyzing the optimal design of mechanisms. For us, the maxmin objective is useful in two respects:

- Motivating our focus on partitional information structures, and
- Simplifying the set of relevant information structures with increasing price paths.

Whether similar simplifications can be obtained for different seller objectives is an open question for future research.

Most of the robust mechanism design literature has considered static settings. Introducing dynamics complicates the characterization of agent behavior, which is essential for understanding the performance of different mechanisms. This difficulty suggests durable-goods pricing as a natural first setting to investigate robust dynamic mechanisms, because a buyer's decision is simply represented by the choice of a stopping time. Nonetheless, one can ask similar questions in other contexts, and we hope the general question of how to design mechanisms under information arrival will receive more attention in the future.

A. PROOFS FOR THE MAIN MODEL

We first define the pressed distribution G in cases where F need not be continuous.

DEFINITION 1'. *Given a percentile $\alpha \in (0, 1]$, define $g(\alpha)$ to be the expected value of the lowest α-percentile of the distribution F. In case F is a continuous distribution, $g(\alpha) = \frac{1}{\alpha}\int_0^{F^{-1}(\alpha)} v dF(v)$. In general, g is continuous and weakly increasing.*

Let $\underline{v}$ be the minimum value in the support of F. For $\beta \in (\underline{v}, \mathbb{E}[v]]$, define $G(\beta) = \sup\{\alpha : g(\alpha) \leq \beta\}$. We extend the domain of this inverse function to $\mathbb{R}_+$ by setting $G(\beta) = 0$ for $\beta \leq \underline{v}$ and $G(\beta) = 1$ for $\beta > \mathbb{E}[v]$.[26]

We now provide proofs of the results for the main model, in the order in which they appeared.

A.1. Proof of Proposition 1

Given a realized price p, minimum profit occurs when there is maximum probability of signals that lead the buyer to have posterior expectation $\leq p$. First consider the information structure $\mathcal{I}$ that tells the buyer whether her value is in the lowest $G(p)$-percentile or above. By definition of G, the buyer's expectation is exactly p upon learning the former. This shows that, under $\mathcal{I}$, the buyer's expected value is $\leq p$ with probability $G(p)$.

Now we show that $G(p)$ cannot be improved upon. To see this, note that it is without loss of generality to consider information structures which recommend the buyer to "buy" or "not buy." Nature chooses an information structure that minimizes the probability of "buy." By Lemma 1 in Kolotilin (2015), this minimum is achieved by a partitional information structure, namely by recommending "buy" for $v > \alpha$ and "not buy" for $v \leq \alpha$. Since the buyer's expected value given $v \leq \alpha$ cannot be greater than p, we have $\alpha \leq F^{-1}(G(p))$. It is then easy to see that the particular information structure $\mathcal{I}$ above is the worst case.

Thus, for any realized price p, the seller's minimum profit is $p(1 - G(p))$. The proposition follows from the seller optimizing over p.

A.2. Proof of Proposition 2

In the main text we showed that for any deterministic price path, nature can choose an information structure that holds profit down to Π^* or lower. Here we extend the argument to any randomized pricing strategy $\sigma \in \Delta(P^T)$. For clarity, the proof will be broken down into three steps.

[26] If F does not have a mass point at $\underline{v}$, $g(\alpha)$ is strictly increasing and $G(\beta)$ is its inverse function which increases continuously. If instead $F(\underline{v}) = m > 0$, then $g(\alpha) = \underline{v}$ for $\alpha \leq m$ and it is strictly increasing for $\alpha > m$. In that case $G(\beta) = 0$ for $\beta \leq \underline{v}$, after which it jumps to m and increases continuously to 1.

Step 1: Cutoff values and information structure. To begin, we define a set of cutoff values. In each period t, given previous and current prices $p_1, \ldots, p_t$, a buyer who knows her value to be v prefers to buy in the current period if and only if

$$v - p_t \geq \max_{\tau \geq t+1} \mathbb{E}[\delta^{\tau-t} \cdot (v - p_\tau)] \tag{7}$$

where the RHS maximizes over all stopping times that stop in the future. It is easily seen that there exists a unique value v_t such that the above inequality holds if and only if $v \geq v_t$.[27] Thus, v_t is defined by the equation

$$v_t - p_t = \max_{\tau \geq t+1} \mathbb{E}[\delta^{\tau-t} \cdot (v_t - p_\tau)] \tag{8}$$

and it is a random variable that depends on realized prices p^t and the expected distribution of future prices $\sigma(\cdot \mid p^t)$.

Next, let us define for each $t \geq 1$

$$w_t = \min\{v_1, v_2, \ldots, v_t\} = \min\{w_{t-1}, v_t\}. \tag{9}$$

For notational convenience, let $w_0 = \infty$ and $w_\infty = 0$. w_t is also a random variable, and it is decreasing over time.

Consider the following information structure $\mathcal{I}$. In each period t, the buyer is told whether or not her value is in the lowest $G(w_t)$-percentile. Providing this information requires nature to know w_t, which depends only on the realized prices and the seller's (future) pricing strategy.

Step 2: Buyer behavior. The following lemma describes the buyer's optimal stopping decision in response to σ and $\mathcal{I}$:

Lemma 1': *For any pricing strategy σ, let the information structure $\mathcal{I}$ be constructed as above. Then the buyer finds it optimal to follow nature's recommendation: she buys when told her value is above the $G(w_t)$-percentile, and she waits otherwise.*

Proof of Lemma 1'. Suppose period t is the first time that the buyer learns her value is above the $G(w_t)$-percentile. Then in particular, $w_t < w_{t-1}$, which implies $w_t = v_t$ by (9). Given this signal, she knows she will receive no more information in the future (because w_t decreases over time). She also knows her value is above the $G(w_t)$-percentile, which is greater than $w_t = v_t$, the average value below that percentile. By the definition of v_t, such a buyer optimally buys in period t.

On the other hand, suppose that in some period t the buyer learns her value is below the $G(w_t)$-percentile. Since w_t decreases over time, this signal contains more information than all

[27] This follows by observing that both sides of the inequality are strictly increasing in v, but the LHS increases faster.

previous signals. By the definition of the pressed distribution G, this buyer's expected value is $w_t \leq v_t$. Such a buyer prefers to delay her purchase even without additional information in the future; the promise of future information does not change the conclusion. ■

Step 3: Profit decomposition. By this lemma, the buyer whose true value belongs to the percentile range $(G(w_{t-1}), G(w_t)]$ will buy in period t. Thus, the seller's expected discounted profit can be computed as

$$\Pi = \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} \cdot (G(w_{t-1}) - G(w_t)) \cdot p_t\right].$$

We rely on a technical result to simplify the above expression:

Lemma 4. *Suppose $w_t = v_t \leq w_{t-1}$ in some period t. Then*

$$p_t = \mathbb{E}\left[\sum_{s=t}^{T-1}(1-\delta)\delta^{s-t}w_s + \delta^{T-t}w_T \mid p^t\right] \tag{10}$$

which is a discounted sum of current and expected future cutoffs.

Using Lemma 4, we can rewrite the profit as

$$\begin{aligned}
\Pi &= \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} \cdot (G(w_{t-1}) - G(w_t)) \cdot \mathbb{E}\left[\sum_{s=t}^{T-1}(1-\delta)\delta^{s-t}w_s + \delta^{T-t}w_T \mid p^t\right]\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} \cdot (G(w_{t-1}) - G(w_t)) \cdot \left(\sum_{s=t}^{T-1}(1-\delta)\delta^{s-t}w_s + \delta^{T-t}w_T\right)\right] \\
&= \mathbb{E}\left[\sum_{s=1}^{T-1}(1-\delta)\delta^{s-1}w_s(1-G(w_s)) + \delta^{T-1}w_T(1-G(w_T))\right] \\
&\leq \Pi^*.
\end{aligned} \tag{11}$$

The second line uses the law of iterated expectations, as well as the fact that w_{t-1} and w_t only depend on the realized prices p^t. The next line follows from interchanging the order of summation, and the last inequality is because $w_s(1-G(w_s)) \leq \Pi^*$ holds for every w_s.

To complete the proof of Proposition 2, it only remains to show Lemma 4.

Proof of Lemma 4. We assume that T is finite.[28] We will prove the result by induction on $T-t$,

[28]The infinite-horizon version can be proved by using finite-horizon approximations and applying the Monotone Convergence Theorem. We omit the technical details.

where the base case $t = T$ follows from $w_T = v_T = p_T$. For $t < T$, from (8) we can find an optimal stopping time $\tau \geq t+1$ such that

$$v_t - p_t = \mathbb{E}[\delta^{\tau-t} \cdot (v_t - p_\tau)]$$

which can be rewritten as

$$p_t = \mathbb{E}[(1-\delta^{\tau-t})v_t + \delta^{\tau-t}p_\tau]. \tag{12}$$

We claim that in any period s with $t < s < \tau$, $v_s \geq v_t$ so that $w_s = w_t = v_t$ by (9); while in period τ, $v_\tau \leq v_t$ and $w_\tau = v_\tau \leq w_{\tau-1}$. In fact, if $s < \tau$, then the optimal stopping time τ suggests that the buyer with value v_t weakly prefers to wait than to buy in period s. Thus by definition of v_s, it must be true that $v_s \geq v_t$. On the other hand, in period τ the buyer with value v_t weakly prefers to buy immediately, and so $v_\tau \leq v_t$.

By these observations, if $\tau = \infty$ (meaning the buyer never buys), we have

$$(1-\delta^{\tau-t})v_t + \delta^{\tau-t}p_\tau = v_t = \sum_{s=t}^{T-1}(1-\delta)\delta^{s-t}w_s + \delta^{T-t}w_T.$$

If $\tau \leq T$, we apply inductive hypothesis to p_τ and obtain

$$(1-\delta^{\tau-t})v_t + \delta^{\tau-t}p_\tau = \sum_{s=t}^{\tau-1}(1-\delta)\delta^{s-t}w_s + \mathbb{E}\left[\sum_{s=\tau}^{T-1}(1-\delta)\delta^{s-t}w_s + \delta^{T-t}w_T \mid p^\tau\right].$$

Plugging the above two expressions into (12) proves the lemma. ■

A.3. Proof of Lemma 2

Fix a dynamic information structure $\mathcal{I}$ and an optimal stopping time τ of the buyer. Because prices are deterministic, the distribution of signal s_t in period t only depends on realized signals (but not prices). Analogously, we can think about the stopping time τ as depending only on past and current signal realizations.

As discussed in the main text, we will construct another information structure $\mathcal{I}'$ which only reveals information in the first period, and which weakly reduces the seller's profit. Consider a signal set $S = \{\overline{s}, \underline{s}\}$, corresponding to the recommendation of "buy" and "not buy", respectively. To specify the distribution of these signals conditional on v, let nature draw signals $s_1, s_2, \cdots$ according to the original information structure $\mathcal{I}$ (and conditional on v). If, along this sequence of realized signals, the stopping time τ results in buying the object, let the buyer receive the signal

$\overline{s}$ with probability $\delta^{\tau-1}$. With complementary probability and when $\tau = \infty$, let her receive the other signal $\underline{s}$. In the alternative information structure $\mathcal{I}'$, nature reveals $\overline{s}$ or $\underline{s}$ in the first period and provides no more information afterwards.

We claim that under $\mathcal{I}'$, the buyer receiving the signal $\underline{s}$ has expected value at most p_1. We actually show something stronger, namely that the buyer has expected value at most p_1 conditional on the signal $\underline{s}$ and *any* realized signal s_1.[29] To prove this, note that since stopping at time τ is weakly better than stopping at time 1, we have

$$\mathbb{E}[v \mid s_1] - p_1 \leq \mathbb{E}^{s_2,\cdots,s_T}\left[\delta^{\tau-1}(\mathbb{E}[v \mid s_1, s_2, \cdots, s_\tau] - p_\tau)\right]. \tag{13}$$

Here and later, the superscripts over the expectation sign highlight the random variables which the expectation is with respect to. In this case they are $s_2, \ldots, s_T$, whose distribution is governed by the original information structure $\mathcal{I}$ and the realized signal s_1.

Since $p_\tau \geq p_1$, simple algebra reduces (13) to the following.

$$\mathbb{E}[v \mid s_1] \leq \mathbb{E}^{s_2,\cdots,s_T}\left[\delta^{\tau-1}\mathbb{E}[v \mid s_1, s_2, \cdots, s_\tau] + (1-\delta^{\tau-1})p_1\right]. \tag{14}$$

Doob's Optional Sampling Theorem says that $\mathbb{E}[v \mid s_1] = \mathbb{E}^{s_2,\cdots,s_T}\left[\mathbb{E}[v \mid s_1, s_2, \cdots, s_\tau]\right]$. Thus we derive the inequality:

$$p_1 \geq \frac{\mathbb{E}^{s_2,\cdots,s_T}[(1-\delta^{\tau-1}) \cdot \mathbb{E}[v \mid s_1, s_2, \cdots, s_\tau]]}{\mathbb{E}^{s_2,\cdots,s_T}[1-\delta^{\tau-1}]}. \tag{15}$$

The denominator $\mathbb{E}^{s_2,\cdots,s_T}[1-\delta^{\tau-1}]$ can be rewritten as $\mathbb{E}^{s_2,\cdots,s_T}\left[\mathbb{P}(\underline{s} \mid s_1, s_2, \ldots, s_T)\right]$, which is the probability of $\underline{s}$ given s_1. Because τ is a stopping time, the numerator in (15) can be rewritten as

$$\mathbb{E}^{s_2,\cdots,s_T}\left[(1-\delta^{\tau-1}) \cdot \mathbb{E}[v \mid s_1, s_2, \cdots, s_T]\right]$$

which can be further rewritten as

$$\mathbb{E}^{s_2,\cdots,s_T}\left[(1-\delta^{\tau-1}) \cdot \mathbb{E}[v \mid s_1, s_2, \cdots, s_T, \underline{s}]\right]$$

because $\underline{s}$ does not provide more information about v beyond $s_1, \ldots, s_T$.

[29] Technically we only consider those s_1 such that $\underline{s}$ occurs with positive probability given s_1.

With these, (15) states that

$$p_1 \geq \frac{\mathbb{E}^{s_2,\cdots,s_T}\left[\mathbb{P}(\underline{s} \mid s_1, s_2, \ldots, s_T) \cdot \mathbb{E}[v \mid s_1, s_2, \cdots, s_T, \underline{s}]\right]}{\mathbb{E}^{s_2,\cdots,s_T}\left[\mathbb{P}(\underline{s} \mid s_1, s_2, \ldots, s_T)\right]} = \mathbb{E}[v|s_1, \underline{s}] \tag{16}$$

just as we claimed.

Thus, under the information structure $\mathcal{I}'$ constructed above, a buyer who receives the signal $\underline{s}$ has expected value at most p_1, which is also less than any future price. Since information only arrives in the first period, all sale happens in the first period to the buyer with the signal $\overline{s}$. The probability of sale is at most $\mathbb{E}[\delta^{\tau-1}]$, and the seller's profit is at most $\mathbb{E}[\delta^{\tau-1}] \cdot p_1$. This is no more than $\mathbb{E}[\delta^{\tau-1} \cdot p_\tau]$, the discounted profit under the original dynamic information structure. We have thus proved that with a deterministic and non-decreasing price path, the seller's profit is at least what he would obtain by selling only once at the price p_1. Taking $p_1 = p^*$ proves the lemma.

A.4. Proof of Theorem 1

By Lemma 2, a constant price path p^* delivers expected un-discounted profit Π^* from each arriving buyer. This matches the upper bound given by Proposition 2 and shows that always charging p^* is optimal. Moreover, suppose p^* is unique in the one-period problem, then from (11) we see that the seller's profit from the first buyer equals Π^* only if $w_s = p^*$ almost surely. This together with Lemma 4 implies $p_1 = p^*$ with probability 1. Similar consideration for later buyers shows that the seller must always charge p^* to achieve the maxmin profit. Hence the theorem.

B. PROOFS FOR THE ALTERNATIVE TIMING MODEL

In this appendix, we first review the solution to the one-period model where information cannot depend on realized price. The analysis follows Du (2018), although we will represent his exponential mechanism as a random price mechanism. After listing several useful properties of Du's mechanism, we will present the proofs of Theorem 2 and Theorem 3. We conclude with the proof of Claim 1, which turns out to build on Theorem 3.

B.1. Properties of Du's Mechanism

For the one-period model, Du (2018) constructs a mechanism that guarantees profit Π_{RSD} regardless of the buyer's information structure. By viewing interim allocation probabilities as a distribution function, we can equivalently implement Du's mechanism as a random price with the following c.d.f.:

$$D(x) = \begin{cases} 0 & x \in [0, W) \\ \frac{\log \frac{x}{W}}{\log \frac{S}{W}} & x \in [W, S) \\ 1 & x \in [S, 1] \end{cases} \tag{17}$$

Recall that W and B are parameters for the Roesler-Szentes information structure (see (4) in Section 7). In the above we have an additional parameter S, which is characterized by $S \in [W, B]$ and

$$\int_0^S F_W^B(v)\, dv = \int_0^S F(v)\, dv \tag{18}$$

where F_W^B is the Roesler-Szentes worst-case information structure. To explain where S comes from, note that the LHS in (18) must not exceed the RHS *for all* S because F is a mean-preserving spread of F_W^B. But when W is smallest possible, such a constraint must bind at some S.

The following observations will be crucial for Claim 1. Since the constraint $\int_0^x F_W^B(v)\, dv \leq \int_0^x F(v)\, dv$ binds at $x = S$, the first order condition gives $F_W^B(S) = F(S)$. This implies that not only F is a mean-preserving spread of F_W^B, but the truncated distribution of F conditional on $v \leq S$ is also a mean-preserving spread of the corresponding truncation of F_W^B. In other words:

Remark 1. *The Roesler-Szentes information structure has the property that any buyer with true value $v \leq S$ has posterior expected value at most S, while any buyer with true value $v > S$ expects her value to be greater than S upon receiving the signal.*

For completeness, we include a quick proof that the random price $p \sim D$ guarantees profit $W = \Pi_{RSD}$. Consider the one-period model in which nature chooses a distribution $\tilde{F}$ of the buyer's posterior expected values. Then the seller's profit is

$$\begin{aligned} \Pi = \int_W^S p(1 - \tilde{F}(p))\, dD(p) &= \frac{1}{\log \frac{S}{W}} \int_W^S (1 - \tilde{F}(p))\, dp \geq \frac{1}{\log \frac{S}{W}} \left(S - W - \int_0^S \tilde{F}(p)\, dp \right) \\ &\geq \frac{1}{\log \frac{S}{W}} \left(S - W - \int_0^S F(p)\, dp \right) = \frac{1}{\log \frac{S}{W}} \left(S - W - \int_0^S F_W^B(p)\, dp \right) = W. \end{aligned}$$

The penultimate equality uses (18) and the last one uses (4).

B.2. Proof of Theorem 2

Consider a constant price p randomly drawn according to Du's distribution $D(\cdot)$ defined above. We will show the seller's discounted expected profit is at least $W = \Pi_{RSD}$.

By assumption, each buyer's expected value follows a martingale process $v_1, v_2, \dots$ that is autonomous (independent of the realized p). As mentioned in the main text, we define a sequence of cutoff prices adapted to the v-process:

$$v_t - r_t = \max_{\tau > t} \mathbb{E}[\delta^{\tau - t}(v_\tau - r_t)]$$

and then

$$q_t = \max\{r_1, \dots, r_t\}.$$

This is exactly dual to the definition of cutoff values, and whenever $q_t = r_t \geq q_{t-1}$, we have (see Lemma 4):

$$v_t = \mathbb{E}\left[\sum_{s \geq t} (1-\delta)\delta^{s-t} q_s \mid v_1, \dots, v_t\right].$$

With foresight, we define the following one-period profit function

$$\pi(y) = \min\{(y - W)^+, S - W\}.$$

That is, $\pi(y) = 0$ for $y \leq W$, $\pi(y) = y - W$ for $y \in [W, S]$ and $\pi(y) = S - W$ for $y \geq S$.

Since purchase occurs in period t precisely when the random price p belongs to $[q_{t-1}, q_t)$, we can compute total profit to be

$$\begin{aligned}
\Pi &= \mathbb{E}\left[\sum_{t \geq 1} \delta^{t-1} \int_{q_{t-1}}^{q_t} p \, dD(p)\right] \\
&= \frac{1}{\log \frac{S}{W}} \cdot \mathbb{E}\left[\sum_{t \geq 1} \delta^{t-1} (\pi(q_t) - \pi(q_{t-1}))\right] \\
&= \frac{1}{\log \frac{S}{W}} \cdot \mathbb{E}\left[\sum_{t \geq 1} (1-\delta)\delta^{t-1} \pi(q_t)\right]
\end{aligned}$$

To facilitate further computation, we introduce a modified version of the function π:

$$\hat{\pi}(y) = \min\{y - W, S - W\} = y - W - (y - S)^+$$

Intuitively, π coincides with $\hat{\pi}$ when $y \geq W$, and is strictly smaller otherwise. Then

$$\begin{aligned}
\log \frac{S}{W} \cdot \Pi &= \mathbb{E}\left[\sum_{t \geq 1}(1-\delta)\delta^{t-1}\pi(q_t)\right] \\
&\geq \mathbb{E}\left[\sum_{t \geq 1}(1-\delta)\delta^{t-1}\hat{\pi}(q_t)\right] \\
&= \mathbb{E}\left[\sum_{t \geq 1}(1-\delta)\delta^{t-1}(q_t - W - (q_t - S)^+)\right] \\
&= v_0 - W - \mathbb{E}\left[\sum_{t \geq 1}(1-\delta)\delta^{t-1}(q_t - S)^+\right]
\end{aligned}$$

where we use the fact that the ex-ante expected value v_0 is a discounted sum of cutoff prices.

Let γ be a stopping time adapted to the v-process such that q_γ first exceeds S. Then we can continue the above computation as follows:

$$\begin{aligned}
\log \frac{S}{W} \cdot \Pi &\geq v_0 - W - \mathbb{E}\left[\sum_{t \geq 1}(1-\delta)\delta^{t-1}(q_t - S)^+\right] \\
&= v_0 - W - \mathbb{E}\left[\delta^{\gamma-1}\sum_{t \geq \gamma}(1-\delta)\delta^{t-\gamma}(q_t - S)\right] \\
&= v_0 - W - \mathbb{E}\left[\delta^{\gamma-1}(v_\gamma - S)\right] \\
&\geq v_0 - W - \mathbb{E}[(v_\infty - S)^+] \\
&= \mathbb{E}[v_\infty - W - (v_\infty - S)^+] \\
&= \mathbb{E}[\hat{\pi}(v_\infty)].
\end{aligned}$$

The third line uses the fact that v_γ is a discounted sum of future cutoff prices, which holds because $q_\gamma \geq S > q_{\gamma-1}$. To show the next inequality, observe that if γ is finite, then $v_\gamma - S \leq \mathbb{E}[(v_\infty - S)^+ \mid v_1, \ldots, v_\gamma]$ by convexity. And if γ is infinite, then $\delta^{\gamma-1}(v_\gamma - S) = 0 \leq (v_\infty - S)^+$.[30]

[30] We highlight that $\hat{\pi}$ is a concave function, whereas the original π is not concave. It is exactly this concavity that allows us to show $\mathbb{E}\left[\sum_{t\geq 1}(1-\delta)\delta^{t-1}\hat{\pi}(q_t)\right] \geq \mathbb{E}[\hat{\pi}(v_\infty)]$, which essentially says that the (limit) distribution of posterior expected values v_∞ is more dispersed than the distribution of cutoff prices q_t. Thus, replacing π with $\hat{\pi}$ is key to this proof.

Let $\tilde{F}$ denote the distribution of v_∞. We have

$$\begin{aligned}
\log \frac{S}{W} \cdot \Pi &\geq \mathbb{E}[\hat{\pi}(v_\infty)] \\
&= \int_0^S (v - W) d\tilde{F}(v) + (S - W)(1 - \tilde{F}(S)) \\
&= \tilde{F}(S)(S - W) - \int_0^S \tilde{F}(v) dv + (S - W)(1 - \tilde{F}(S)) \\
&= S - W - \int_0^S \tilde{F}(v) dv \\
&\geq S - W - \int_0^S F(v) dv \\
&= \log \frac{S}{W} \cdot W
\end{aligned}$$

The inequality follows from F being a mean preserving spread of $\tilde{F}$, and the last equality follows from (4). Hence $\Pi \geq W$ as desired. ■

B.3. Proof of Theorem 3

We will construct a dynamic pricing strategy that guarantees Π_{RSD} from each buyer. The construction relies on the following lemma regarding the outcome-equivalence between static and dynamic pricing strategies:

Lemma 5. *Fix any continuous distribution function $D(\cdot)$,[31] any horizon T and any discount factor $\delta \in (0, 1)$. There exists a distribution of prices $\sigma \in \Delta(p^T)$ such that if a buyer arrives in period t and knows her value to be v, then her discounted probability of purchasing the object (discounted to period t) is equal to $D(v)$.*

In words, for any static pricing strategy there is a dynamic pricing strategy which does not condition on buyers' arrival times, but which results in the same discounted purchase probabilities for *every type of each arriving buyer*. As a consequence, a seller using strategy σ obtains the same profit from any buyer as if he sells only once to this buyer at a random price drawn from D. This is true whenever the buyer's value distribution is determined upon arrival and fixed over time, which is what we assume for the current theorem. Since Du's static mechanism guarantees profit Π_{RSD} from each buyer, Theorem 3 will follow once we prove the lemma.

[31] Note that Du's distribution $D(\cdot)$ is continuous except when it is a point-mass on W; in that exceptional case we have $\Pi_{RSD} = \Pi^*$, and Theorem 3 follows from Theorem 1.

Proof of Lemma 5. We will first prove the result for $T = 2$, then generalize to all finite T and lastly discuss $T = \infty$.

Step 1: The case of two periods. In the second period, regardless of realized p_1 the seller should charge a random price drawn from D. This achieves the desired allocation probabilities for the second buyer.

Consider the first buyer. For any price p_1 in the first period, define v_1 as the cutoff indifferent between buying at price p_1 or waiting till the next period and facing the random price drawn from D. That is,

$$v_1 - p_1 = \delta \cdot \mathbb{E}^{p_2 \sim D} \left[\max\{v_1 - p_2, 0\}\right]. \tag{19}$$

As p_1 varies according to the seller's pricing strategy σ, v_1 is a random variable. As in the proof of Proposition 2, we define $w_1 = v_1$ and $w_2 = \min\{v_1, p_2\}$, where p_2 is independently drawn according to D.

If the buyer has value $x \geq w_1$, she buys in the first period. Otherwise if she has value $w_1 > x \geq w_2$, she buys in the second period. The discounted purchase probability of such a buyer is thus

$$\mathbb{P}^{w_1}[x \geq w_1] + \delta \cdot \mathbb{P}^{w_1,w_2}[w_1 > x \geq w_2] = (1-\delta) \cdot \mathbb{P}^{w_1}[x \geq w_1] + \delta \cdot \mathbb{P}^{w_2}[x \geq w_2].$$

Let w be the random variable that satisfies $w = w_1$ (or w_2) with probability $1 - \delta$ (or δ), then the seller seeks to ensure that w is distributed according to D.

Suppose H is the c.d.f. of v_1. Since $w_1 = v_1$ and $w_2 = \min\{v_1, p_2\}$, the probability that w is *greater than* x is given by $(1-\delta)(1-H(x)) + \delta(1-H(x))(1-D(x))$.[32] This has to be equal to $1 - D(x)$, which implies

$$1 - H(x) = \frac{1 - D(x)}{1 - \delta D(x)}. \tag{20}$$

We are left with the task of finding a first-period price distribution under which $v_1 \sim H$. This can be done because the random variables v_1 and p_1 are in a one-to-one relation (see (19)). The lemma thus holds for $T = 2$.

Before proceeding, we remark that (20) implies the distribution H has the same support as D. However, (19) suggests that when v_1 achieves the maximum of this support, p_1 is in general strictly smaller than v_1 (unless the support is a singleton point, a degenerate case). On the other hand, the minimum price p_1 is indeed equal to the minimum of the support of D; when D is Du's distribution, this minimum is exactly W.

[32] $1 - H(x)$ is the probability that $w_1 > x$, and $(1 - H(x))(1 - D(x))$ is the probability that $w_2 > x$.

Step 2: Extension to finite T. We conjecture a pricing strategy σ that is independent across periods: $d\sigma(p_1, \dots, p_T) = d\sigma_1(p_1) \times \cdots \times d\sigma_T(p_T)$, where we interpret each σ_t as a distribution. Define the cutoff values $v_1, \dots, v_T$ as in (8). Note that due to independence, v_t only depends on current price p_t but not on previous prices.

Consider a buyer who arrives in period t. We can generalize the previous arguments and show that if she knows her value to be x, then her discounted purchase probability is $\mathbb{P}[w^{(t)} \le x]$. The random variable $w^{(t)}$ is described as follows: for $t \le s \le T-1$, $w^{(t)} = \min\{v_t, v_{t+1}, \dots, v_s\}$ with probability $(1-\delta)\delta^{s-t}$; and with remaining probability δ^{T-t}, $w^{(t)} = \min\{v_t, v_{t+1}, \dots, v_T\}$.

The result of the lemma requires each $w^{(t)}$ to be distributed according to D. Simple calculation shows this is the case if $v_T \sim D$ and $v_1, \dots, v_{T-1} \sim H$ (since v_t depends only on p_t, they are independent random variables).[33] We can then solve for the price distributions $\sigma_1, \dots, \sigma_T$ by backward induction: σ_T must be D, and once the prices in period $t+1, \dots, T$ are determined, there is a one-to-one relation between p_t and v_t by (8). Thus, the distribution of p_t is uniquely pinned down by the desired distribution of v_t.

Step 3: The infinite horizon case. If $T = \infty$, we look for price distributions $\sigma_1, \sigma_2, \dots$ such that $v_1, v_2, \dots \sim H$. We conjecture a stationary σ_t. Recall that the cutoff v_1 is defined by

$$v_1 - p_1 = \max_{\tau \ge 2} \mathbb{E}\left[\delta^{\tau-1}(v_1 - p_\tau)\right]. \tag{21}$$

The stopping problem on the RHS is stationary. Thus when $p_2 < p_1$ the buyer stops in period 2 and receives $v_1 - p_2$; otherwise she continues and receives $v_1 - p_1$. (21) thus reduces to

$$v_1 - p_1 = \delta \cdot \mathbb{E}^{p_2}\left[\max\{v_1 - p_1, v_1 - p_2\}\right]$$

which can be further simplified to

$$v_1 = p_1 + \frac{\delta}{1-\delta} \cdot \mathbb{E}^{p_2}\left[\max\{p_1 - p_2, 0\}\right]. \tag{22}$$

Let $P(x)$ denote the c.d.f. of p_1 (and of p_2). When $p_1 = x$, (22) implies

$$v_1 = x + \frac{\delta}{1-\delta} \cdot \int_0^x (x-z)\, dP(z) = x + \frac{\delta}{1-\delta} \int_0^x P(z)\, dz.$$

[33] The reason $H(x)$ should be the c.d.f. of v_1 is best understood in the infinite horizon problem (see below). Under stationarity, the buyer with value x buys in period t with probability $H(x)$, conditional on not buying previously. Thus the discounted allocation probability is $\sum_t \delta^{t-1}(1-H(x))^{t-1}H(x)$. Setting this equal to $D(x)$ yields (20).

Thus v_1 has c.d.f. $H(x)$ if and only if

$$P(x) = H\left(x + \frac{\delta}{1-\delta}\int_0^x P(z)\,dz\right). \tag{23}$$

To solve for $P(x)$, we let

$$Q(x) = x + \frac{\delta}{1-\delta}\int_0^x P(z)\,dz; \qquad U(y) = 1 + \frac{\delta}{1-\delta}H(y) = \frac{1}{1-\delta D(y)}. \tag{24}$$

(23) is the differential equation

$$U(Q(x)) = Q'(x). \tag{25}$$

Put $V(y) = \int_0^y (1-\delta D(z))\,dz$, so that $V'(y) = \frac{1}{U(y)}$. Then

$$\frac{\partial V(Q(x))}{\partial x} = V'(Q(x)) \cdot Q'(x) = \frac{Q'(x)}{U(Q(x))} = 1. \tag{26}$$

Inspired by the analysis for finite T, we conjecture that the minimum value of p_1 is W. That is, we conjecture $Q(W) = W$. Since $V(W) = W$, we deduce from (26) that $Q(x)$ is characterized by

$$V(Q(x)) = x \;\text{ with }\; V(y) = \int_0^y (1-\delta D(z))\,dz. \tag{27}$$

Since V is strictly increasing, there is a unique solution $Q(x)$ to the above equation, and the corresponding distribution of prices is

$$P(x) = \frac{1-\delta}{\delta} \cdot (Q'(x) - 1). \tag{28}$$

Lemma 5 is proved, and so is Theorem 3. ■

B.4. Proof of Claim 1

The proof is somewhat long, and we will present it in several steps.

Step 1: The information structure. Consider now the model with two periods and one buyer arriving in each period. By providing the Roesler-Szentes information structure to the second buyer, nature can ensure that seller obtains no more than Π_{RSD} from her.

For the *first* buyer, we construct the following dynamic information structure $\mathcal{I}$:

- In the first period, nature provides the Roesler-Szentes information structure. We denote the buyer's posterior expected value by $\tilde{v}$, so as to distinguish from her true value v. Note that $\tilde{v} \sim F_W^B$.

- In the second period, given the realized price p_1 and the buyer's expected value $\tilde{v}$ in the first period, nature reveals the buyer's true value v if $\tilde{v} \geq v_1(p_1)$; otherwise nature provides no additional information. Here the cutoff $v_1(p_1)$ is defined as usual (assuming no information arrives in the second period):

$$v_1 - p_1 = \delta \cdot \mathbb{E}^{p_2 \sim \sigma(\cdot|p_1)} \left[\max\{v_1 - p_2, 0\}\right].$$

Intuitively, nature targets the buyer who prefers to buy in the first period when she does not expect to receive information in the second period. By promising full information to such a buyer in the future, nature potentially delays her purchase and reduces the seller's profit. In what follows we formalize this intuition.

Step 2: Buyer behavior and seller profit. To facilitate the analysis, we consider a simpler information structure $\mathcal{I}'$ in which nature reveals $\tilde{v}$ in the first period but does nothing in the second period. Under $\mathcal{I}'$, the buyer's value distribution F_W^B does not change over time. Thus by Stokey (1979), the seller's profit would at most be Π_{RSD}. We will show that the seller's profit under the *dynamic* information structure $\mathcal{I}$ could only be lower than under $\mathcal{I}'$ (for any pricing strategy), and we also characterize when the comparison is strict.

There are three possibilities. First, if the price p_1 is relatively high so that $\tilde{v} < v_1(p_1)$, then the buyer does not buy in the first period under $\mathcal{I}'$. This is also her optimal decision under $\mathcal{I}$, because she will not receive extra information in the second period. Second, if the price is very low, then under both $\mathcal{I}$ and $\mathcal{I}'$ the buyer buys in the first period. Lastly, for some intermediate prices the buyer buys in the first period under $\mathcal{I}'$ but not under $\mathcal{I}$. The opposite case cannot arise because $\mathcal{I}$ provides more information than $\mathcal{I}'$ in the second period, so the buyer has a stronger incentive to wait.

Thus, when nature provides $\mathcal{I}$ rather than $\mathcal{I}'$, the seller's profit changes only in the last possibility above. Now observe that whenever the buyer chooses to delay purchase, the discounted social surplus decreases by at least $(1-\delta)\tilde{v} \geq (1-\delta)W$. Since the buyer's expected payoff cannot decrease, this means the seller's profit decreases by at least $(1-\delta)W$.

To summarize, we have shown:

Lemma 6. *Consider the information structures $\mathcal{I}$ and $\mathcal{I}'$ constructed above. The seller's profit under $\mathcal{I}'$ is no greater than Π_{RSD}, and his profit under $\mathcal{I}$ is even smaller by at least $(1-\delta)W$ times the* *probability that the buyer delays purchase.*

Step 3: Proof of the claim for σ^D. Let σ^D be the pricing strategy given by Lemma 5, which we recall is robust to information that arrives only upon arrival (for each buyer). Here we argue that if the seller uses σ^D, then his profit from the first buyer is strictly less than Π_{RSD} under the above *dynamic* information structure $\mathcal{I}$. Later we generalize the result to other pricing strategies.

Recall from the proof of Lemma 5 that under σ^D, p_2 is drawn from Du's distribution D, independently of p_1. On the other hand, p_1 is continuously supported on a smaller interval $[W, S_1]$, with $W < S_1 < S$. More precisely, the distribution of p_1 is determined by the condition that $v_1(p_1) \sim H$; see (20).

Suppose the buyer's posterior expected value $\tilde{v}$ (in the first period) belongs to the open interval (W, S). Further suppose that knowing her true value *strictly* improves her expected payoff in the second period (given $p_2 \sim D$). Then, whenever p_1 is smaller than but close to $v_1^{-1}(\tilde{v})$, such a buyer would buy in the first period under $\mathcal{I}'$ but delay purchase under $\mathcal{I}$. Applying Lemma 6, it only remains to find a positive measure of such buyers.

Indeed, from Remark 1 we know that $\tilde{v} < S$ implies the true value also satisfies $v < S$. Moreover, because we assume $\Pi_{RSD} > \Pi^*$, Proposition 5 in Appendix D.3 gives $W > \underline{v}$. Thus with positive probability, a buyer with $\tilde{v} \in (W, S)$ has true value $v \in (\underline{v}, W)$. For any such buyer, because her expected value $\tilde{v}$ exceeds W, she purchases at some price $p_2 \sim D$ without additional information. But if she were informed that $v < W$, she would not buy at any second-period price p_2 (which is at least W). Hence knowing her true value strictly improves her second-period expected payoff, and we are done with the proof here.

Step 4: Proof for an arbitrary pricing strategy σ. Finally, we turn to prove the claim in its full generality. The argument is as follows (omitting technical details): Suppose for contradiction that some pricing strategy σ guarantees profit almost Π_{RSD} from each buyer. Then because D is uniquely optimal in the one-period problem, the distribution of p_2 conditional on p_1 is "close to" D (in the Prokhorov metric) with high probability; otherwise nature could sufficiently damage the seller's profit from the second buyer. Next, we can similarly show that the distribution of $v_1(p_1)$ under σ must be close to H (which is its distribution under σ^D).[34] The rest of the proof proceeds as in Step 3: With positive probability the buyer has true value $v < W$ and posterior expected

[34] If Du's mechanism is not unique in the one-period problem, then we cannot reach these conclusions. In fact, without that technical assumption *our proof* presented here would fail: In an earlier version of this paper we show that if nature provides the specific information structure $\mathcal{I}$, then whenever Du's mechanism is non-unique, the seller *has* a pricing strategy that obtains Π_{RSD} from both buyers. Details are available upon request.

value $\tilde{v} \in (W, S)$. For such a buyer, full information in the second period is strictly valuable, and she delays purchase under the information structure $\mathcal{I}$ relative to $\mathcal{I}'$. Lemma 6 then implies that under $\mathcal{I}$, profit from the first buyer is bounded away from Π_{RSD}. This leads to a contradiction, and the proof of Claim 1 is complete. ■

C. PROOFS FOR THE COMMON VALUE MODEL

C.1. Proof of Lemma 3

Fixing any (public) dynamic information structure $\mathcal{I}$, we will replace it with another information structure $\mathcal{I}'$ that only provides a single public signal in the first period. Moreover, as in the proof of Lemma 2, we will ensure that each arriving buyer has lower discounted purchase probability under the replacement $\mathcal{I}'$. This will show that the seller obtains lower profit under $\mathcal{I}'$.

To do this, consider any possible signal history $s_1, s_2, \dots$ under the original process $\mathcal{I}$. For each arriving buyer a, let τ_a denote her optimal stopping time along this history; that is, the buyer who arrives in period a finds it optimal to purchase in period τ_a given the signal realizations $s_1, \dots, s_{\tau_a}$. Note that we always have $\tau_a \geq a$, and $\tau_{a+1} \geq \tau_a$ with equality whenever $\tau_a > a$. The latter property derives from our assumption that signals are publicly observed.

We define a "critical" subset of buyers $j_1, j_2, \dots$ as follows: To begin, j_1 is the first buyer who delays purchase (i.e. with $\tau_{j_1} > j_1$). Next, j_2 is the first buyer after τ_{j_1} such that $\tau_{j_2} > j_2$. So on and so forth, until every later buyer purchases immediately upon arrival. When that occurs we complete the definition by including a hypothetical buyer $j = T + 1$ into the subset.

As an example, suppose $T = 7$, and buyers' stopping times are $2, 2, 3, 6, 6, 6, 7$. Then buyers $1, 4, 8 (= T + 1)$ are critical. More generally, it is not difficult to show that the critical buyers and *their* stopping times uniquely determine the stopping behavior of all the buyers.

Now we are ready to construct the replacement information structure $\mathcal{I}'$. We assume the signal set is $\{0, 1, \dots, T\}$, where the signal "i" represents nature's recommendation that buyers with $a \leq i$ purchase upon arrival and that other buyers do not purchase. Furthermore, given the original signal history $s_1, s_2, \dots$, we assume that signal i realizes only if $i = j_m - 1$ for some critical buyer j_m. We specify the probability of such a signal to be[35]

$$\delta^{\sum_{k<m} \tau_{j_k} - j_k} \cdot \left(1 - \delta^{\tau_{j_m} - j_m}\right).$$

To interpret, these probabilities are such that $i \geq j_m$ occurs with *conditional* probability

[35] Instead of introducing the critical subset and writing out the signal probabilities in closed form (as done here), one can also prove the result by induction on T and recursively define the signal probabilities.

$\delta^{\tau_{j_m}-j_m}$ conditional on $i \geq j_m - 1$. In other words, the replacement information structure recommends the critical buyer j_m to purchase with conditional probability $\delta^{\tau_{j_m}-j_m}$. This is in line with the proof of Lemma 2 since we "push and discount" nature's recommendation to the buyer's arrival time. Due to conditioning, however, a difference arises here in that $\delta^{\tau_{j_m}-j_m}$ is *not* the probability of receiving a signal $i \geq j_m$ (except for $m = 1$). From the above formula, we see that any critical buyer is recommended to purchase with probability *smaller than* $\delta^{\tau_{j_m}-j_m}$; in fact, this holds also for non-critical buyers.[36]

Recommending lower purchase probabilities serves our goal, which is to show that seller's profit is lower under $\mathcal{I}'$ than under $\mathcal{I}$. Nonetheless, we still have to verify that buyers want to follow nature's recommendation when it comes to *not purchasing* the object.[37] Suppose buyer a receives a signal $i < a$, we need to argue that her expected value is lower than p_a. Since all buyers have the same expectation and prices are increasing over time, it is sufficient to consider $a = i + 1$. Then by definition, a must be a critical buyer j_m.

We will prove a stronger result, that conditional on any realizations $s_1, \ldots, s_{j_m}$ (and on the signal i), expected value is at most p_a. Indeed, once $s_1, \ldots, s_{j_m}$ are fixed, so are the critical buyers before j_m as well as their stopping times. Thus the term $\delta^{\sum_{k<m} \tau_{j_k}-j_k}$ is simply a multiplicative constant in the probability of signal i. This suggests that the expected value conditional on signal i is unchanged if we instead specify the probability of this signal to be $1 - \delta^{\tau_{j_m}-j_m}$. But then we return to the proof of Lemma 2, where the buyer is recommended to not purchase with probability $1 - \delta^{\tau_{j_m}-j_m}$. Hence the result here follows from that proof. ■

C.2. Proof of Theorem 4

From (5) in the main text, the seller's maxmin profit with an increasing price path is

$$(1-\delta)\Pi^C = \min_{\tilde{F}} \sum_{t=1}^{T} (1-\delta)\delta^{t-1} p_t \cdot (1 - \tilde{F}(p_t)).$$

The RHS can be interpreted as the profit in the one-period problem, when the seller charges a random price that is equal to p_t with probability $(1-\delta)\delta^{t-1}$. Thus, as long as the seller chooses $p_1, \ldots, p_T$ such that the distribution of this random price approximates Du's distribution $D(\cdot)$, he can guarantee profit close to Π_{RSD}.

[36] The probability that any buyer a receives a signal $i \geq a$ is $\delta^{\sum_{k \leq m} \tau_{j_k}-j_k}$, where j_m is the last critical buyer up to and including a. Since we always have $\tau_{j_m} - j_m \geq \tau_a - a$, this probability is lower than $\delta^{\tau_a - a}$.

[37] Recall that we also had to prove this for Lemma 2; on the other hand, whether or not the buyers follow the recommendation to purchase does not affect our argument.

To achieve this approximation, we equate the c.d.f.s at the discrete points $p_1, \ldots, p_T$. This leads to prices defined by

$$D(p_t) = 1 - \delta^t,$$

or equivalently

$$p_t = W \cdot (S/W)^{1-\delta^t}$$

As $\delta \to 1$ and $T \to \infty$, these points $p_1 \sim p_T$ are densely distributed on the interval (W, S). Hence their distribution converges to $D(\cdot)$, which proves the theorem. ■

D. OTHER RESULTS

D.1. Uncertainty Leads to Lower Price

We prove here that uncertainty over the information structure leads the seller to choose a lower price than if the buyer knew her value.

Proposition 4. *For any continuous distribution F, let $\hat{p}$ be an optimal monopoly price under known values:*

$$\hat{p} \in \underset{p}{\operatorname{argmax}}\, p(1 - F(p)). \tag{29}$$

Then any maxmin optimal price p^* satisfies $p^* \leq \hat{p}$. Equality holds only if $p^* = \hat{p} = \underline{v}$.

Proof of Proposition 4. It suffices to show that the function $p(1 - G(p))$ strictly decreases when $p > \hat{p}$, until it reaches zero. By taking derivatives, we need to show $G(p) + pG'(p) > 1$ for $p > \hat{p}$ and $G(p) < 1$.

From definition, the lowest $G(p)$-percentile of the distribution F has expected value p. That is,

$$pG(p) = \int_0^{F^{-1}(G(p))} v dF(v), \forall p \in [\underline{v}, \mathbb{E}[v]]. \tag{30}$$

Differentiating both sides with respect to p, we obtain

$$G(p) + pG'(p) = \frac{\partial}{\partial p}(F^{-1}(G(p))) \cdot F^{-1}(G(p)) \cdot F'(F^{-1}(G(p))) = G'(p) \cdot F^{-1}(G(p)). \tag{31}$$

This enables us to write $G'(p)$ in terms of $G(p)$ as follows:

$$G'(p) = \frac{G(p)}{F^{-1}(G(p)) - p}. \tag{32}$$

Thus,

$$G(p) + pG'(p) = \frac{G(p) \cdot F^{-1}(G(p))}{F^{-1}(G(p)) - p}. \tag{33}$$

We need to show that the RHS above is greater than 1, or that $F^{-1}(G(p)) < \frac{p}{1-G(p)}$ whenever $p > \hat{p}$ and $G(p) < 1$. This is equivalent to $G(p) < F(\frac{p}{1-G(p)})$, which in turn is equivalent to

$$\frac{p}{1-G(p)} \cdot \left(1 - F\left(\frac{p}{1-G(p)}\right)\right) < p. \tag{34}$$

From the definition of $\hat{p}$, we see that the LHS above is at most $\hat{p}(1 - F(\hat{p})) \leq \hat{p} < p$, as we claim to show. Moreover, when $\hat{p} > \underline{v}$, the last inequality $\hat{p}(1 - F(\hat{p})) < \hat{p}$ is strict. Tracing back the previous arguments, we see that $G(p) + pG'(p) > 1$ holds even at $p = \hat{p}$. In that case we would have the strict inequality $p^* < \hat{p}$ as desired. ■

D.2. Uniqueness of Du's Mechanism

Recall the random price mechanism from Appendix B.1. In general, there could be more than one point S for which (18) holds. Then there might not be a unique optimal mechanism for the single-period model. Nonetheless, for "generic" distributions F, the point S is indeed unique.[38] The following result verifies that the optimal mechanism is unique whenever S is uniquely defined.[39]

Lemma 7. *There is a unique maxmin-optimal mechanism in the one-period alternative timing model if and only if (18) holds at a unique point S.*

Proof of Lemma 7. We focus on the "if" direction. Suppose S is unique, we need to show any random price that guarantees W must follow Du's distribution D. Let $r(p)$ be the p.d.f. of the random price, then seller's profit is given by

$$\Pi = \int_0^1 p \cdot r(p) \cdot (1 - \tilde{F}(p)) \, dp. \tag{35}$$

Given seller's choice $r(p)$, nature chooses a c.d.f. $\tilde{F}$ to minimize Π. Nature's constraint is that F

[38] The intuition is simple: (18) must bind at some S, but for it to bind at two different points would impose a non-generic constraint on F. We omit the formal proof, which is tangential to the paper.

[39] A sufficient condition is that the c.d.f. $F(x)$ is convex. To show this, note that $F(x) - F_W^B(x) = F(x) + \frac{W}{x} - 1$ is convex, so it has at most two roots $x_0 < x_1$. Since $F(x) > F_W^B(x)$ for $x < x_0$, (18) implies S cannot be the smaller root x_0. Hence S must be the bigger root x_1.

must be a mean-preserving spread of $\tilde{F}$. That is,

$$\int_0^x \tilde{F}(v)\,dv \leq \int_0^x F(v)\,dv,$$

for all $x \in (0,1]$, with equality at $x = 1$.

By Roesler and Szentes (2017), $\tilde{F} = F_W^B$ is a solution to nature's problem. For this solution, the above integral inequality constraint *only* binds at $x = S$. Standard perturbation techniques thus imply that nature cannot improve upon $\tilde{F} = F_W^B$ only if $p \cdot r(p)$ is a constant for $p \in (W, S)$. Indeed, suppose that $p \cdot r(p) > p' \cdot r(p')$ for some $p, p' \in (W, S)$. Then starting with $\tilde{F} = F_W^B$, nature could increase $\tilde{F}$ around p and correspondingly decrease it around p'. The perturbed distribution $\tilde{F}$ is still feasible, but the profit is reduced.

Similarly, $p \cdot r(p)$ must also be a constant on the interval $p \in (S, B)$. Let c be this constant, and suppose nature fix $\tilde{F}$ to be F_W^B on the interval $[0, S]$. Then, on the interval $[S, B]$, it seeks to minimize $c \cdot \int_S^B (1 - \tilde{F}(v))\,dv$ subject to $\int_S^1 (1 - \tilde{F}(v))\,dv = \int_S^1 (1 - F(v))\,dv$. This is equivalent to maximizing $c \cdot \int_B^1 (1 - \tilde{F}(v))\,dv$. Since the choice of $\tilde{F} = F_W^B$ results in 0, we must have $c = 0$ in order for F_W^B to be nature's optimal strategy.

To summarize, we have shown that $r(p)$ must be supported on $[W, S]$ and $p \cdot r(p)$ is a constant. This condition together with $\int_W^S r(p)\,dp = 1$ uniquely pins down $r(p)$, which must be the density function associated with Du's distribution. ■

D.3. Comparison Between Π^* and Π_{RSD}

Here we show that the profit benchmark Π_{RSD} is in general higher than Π^*, and the difference may be significant:

Proposition 5. $\Pi_{RSD} \geq \Pi^*$ *with equality if and only if* $W = \underline{v}$ $(= p^*)$, *where* W *is as defined in the Roesler-Szentes information structure (4). Furthermore, as the distribution* F *varies, the ratio* Π_{RSD}/Π^* *is unbounded.*

Proof of Proposition 5. The inequality $\Pi_{RSD} \geq \Pi^*$ is obvious. Next, recall that $\Pi^* \geq \underline{v}$ (seller can charge $\underline{v}$) and $W = \Pi_{RSD}$. Thus $W = \underline{v}$ implies $\Pi_{RSD} \leq \Pi^*$, and equality must hold.

Conversely suppose $\Pi_{RSD} = \Pi^*$, then $W = p^*(1 - G(p^*))$. This implies $p^* \geq W$. Consider a seller who charges price p^* against the Roesler-Szentes information structure F_W^B. By the unit elasticity of demand property, this seller's profit is either W (when $p^* < B$) or 0. We have shown in our one-period model that the seller can guarantee Π^* with a price of p^*. Thus the seller's profit must be W when he charges p^* and nature chooses the Roesler-Szentes information structure. Since $W = \Pi^*$ by assumption, the Roesler-Szentes information structure is a worst-case information structure for the price p^*. This yields $W \geq p^*$, because a worst-case information structure cannot include any signal that leads to a posterior expected value strictly less than p^*. We conclude $p^* = W = p^*(1 - G(p^*))$, from which it follows that $G(p^*) = 0$ and $p^* = \underline{v}$. Thus $W = \underline{v}$ must hold.

To study the ratio Π_{RSD}/Π^*, we restrict attention to a very simple class of distributions F: with probability λ, the buyer's true value is 1; otherwise her value is 0. The optimal price in the known-value case is $\hat{p} = 1$, and the corresponding profit is $\hat{\Pi} = \lambda$. In our main model, the maxmin optimal price p^* solves

$$p^* \in \operatorname*{argmax}_p p(1 - G(p)) = \operatorname*{argmax}_{0 \leq p \leq \lambda} p \cdot \frac{\lambda - p}{1 - p}$$

Simple algebra gives $p^* = 1 - \sqrt{1 - \lambda}$, and $\Pi^* = (1 - \sqrt{1 - \lambda})^2$ which is roughly $\frac{\lambda^2}{4}$ for small λ.

Because the distribution F has two-point support, it is clear that nature can induce any $\tilde{F}$ supported on $[0, 1]$ with mean λ as the distribution of posterior expected values. Thus the Roesler-Szentes information structure involves the smallest W such that F_W^B has mean λ for some $B \leq 1$. From (4), we compute that the mean of F_W^B is $W \log B - W \log W + W$. We look for the smallest W such that $\log B = \frac{\lambda}{W} + \log W - 1$ is non-positive. It follows that W is the smallest positive root of the equation

$$\frac{\lambda}{W} + \log W = 1.$$

For λ small, we have the approximation $\Pi_{RSD} = W \approx \frac{\lambda}{|\log \lambda|}$. Thus both ratios $\hat{\Pi}/\Pi_{RSD}$ and Π_{RSD}/Π^* are unbounded.[40] ■

D.4. Alternative Interpretation of Π^*

In this appendix, we consider a static information acquisition game that also yields our solution (to our main model). The motivation borrows from Roesler and Szentes (2017), so we begin by reviewing their result.

Roesler and Szentes (2017) consider a game with the following timing: The buyer first chooses an information structure $\mathcal{I} : \mathbb{R}_+ \to \Delta(S)$. The seller then chooses a price $p \in \mathbb{R}$. Finally, the buyer observes her signal and decides whether or not to purchase the object. Those authors show that in order to maximize payoff, the buyer acquires information according to F_W^B. This turns out to simultaneously minimize the seller's profit.

[40] We conjecture that these profit ratios become bounded under certain regularity conditions on F.

Recall that *our* one-period model differs from Roesler and Szentes (2017) in that we allow nature to provide information depending on the realized price. Inspired by this difference, we modify the above information acquisition game so that the buyer can acquire information based on the price. That is, we maintain the same timing as above, except that the buyer chooses a price-dependent information structure $\mathcal{I} : V \times P \to \Delta(S)$.[41]

We characterize the outcome of this game in the following result:

Proposition 6. *Consider the above information acquisition game where the buyer (or benevolent third party) chooses a price-dependent information structure. In any Nash equilibrium of this game, the seller's profit is* Π^* *and the buyer's expected payoff is* $\mathbb{E}[v] - \Pi^*$.[42]

Proof of Proposition 6. For each price p, let $\mathcal{I}^*(p)$ be the corresponding worst-case partitional information structure in our main model. We first construct a (subgame-perfect) equilibrium as follows: On the equilibrium path, the buyer chooses to acquire no information if $p = \Pi^*$, but for any other price he acquires information according to $\mathcal{I}^*(p)$; the seller chooses $p = \Pi^*$. Off the equilibrium path, the buyer chooses a different information structure, and the seller best responds with some price.

To see this is an equilibrium, observe that on path, trade occurs with probability 1 because $\Pi^* < \mathbb{E}[v]$ whenever F is non-degenerate. Hence seller's profit is Π^* and buyer's payoff is $\mathbb{E}[v] - \Pi^*$, sharing all the surplus. By the definition of Π^*, choosing $p = \Pi^*$ is the seller's best response on the equilibrium path. It remains to check that the buyer cannot profitably deviate. Indeed, regardless of the buyer's choice of information structure, the seller can always set price to be p^* and guarantee profit Π^*. Since the seller best responds, her actual profit must be higher. But total surplus cannot exceed $\mathbb{E}[v]$, which implies that buyer's payoff is at most $\mathbb{E}[v] - \Pi^*$. This verifies our construction.

Since this is a sequential-move game, the same argument shows that buyer's payoff must be $\mathbb{E}[v] - \Pi^*$ in *every* equilibrium. Again because total surplus is bounded by $\mathbb{E}[v]$, seller's profit cannot exceed Π^*. But we have argued that he can guarantee Π^*, so this must be his profit level in every equilibrium. Hence the proposition. ■

Note that the same argument works for an arbitrary horizon. That is, suppose the buyer chooses a (price-dependent) *dynamic* information structure to maximize his payoff, whereas the

[41]We implicitly require the buyer to *commit to* acquiring information according to $\mathcal{I}$ after the price is realized. It may be difficult to assume that the buyer literally does this in practice. However, such information may be provided by a third party whose objective is to help the buyer (rather than directly hurt the seller).

[42]Similar to Roesler and Szentes (2017), trade occurs with probability 1 in equilibrium. Unlike in that paper, however, for us the buyer's payoff is higher in this game than in the worst-case scenario for the seller.

seller responds with a pricing strategy. Then in every equilibrium of this game, buyer receives $\mathbb{E}[v] - \Pi^*$ and seller obtains Π^*.

D.5. Known Information Arrival Process

This appendix supplies two simple examples, both with one buyer and two periods, where the information arrival process is *known* to the seller. We show that optimal prices in such a problem can be increasing or decreasing over time. These examples highlight the difficulty in obtaining clear predictions without using the robustness approach. Buyer learning does not by itself predict the optimality of constant prices; seller uncertainty (in our main model) is also important for this conclusion.

Example 3. *In the first example, consider any prior distribution F. Suppose nature reveals nothing in the first period and everything in the second period. Then the seller can and should sell only in the first period to obtain the full surplus. For high δ, any pair of optimal prices must satisfy $p_1 < p_2$ in order to prevent the buyer from waiting.*

Example 4. *In the second example, suppose the buyer is one of two types, L or H, with equal probabilities. Type H buyer has value equal to 1 with certainty. Type L buyer has value $\frac{2}{3}$ with probability $3/4$ and value 0 with probability $1/4$. Nature tells the buyer her type in the first period and reveals her value in the second period; note that this is a partitional information arrival process. It is not difficult to show that the (uniquely) optimal prices are $p_1 = 1 - \frac{\delta}{3}$ and $p_2 = \frac{2}{3}$—the choice of p_1 is such that the H type buyer is exactly indifferent between purchasing in either period. In particular, the resulting profit $\frac{1}{2} + \frac{\delta}{12}$ is higher than selling only in the first period (profit $\frac{1}{2}$) or selling only in the second period (profit $\frac{7\delta}{12}$). Here, $p_1 > p_2$ and the seller benefits from intertemporal price discrimination.*

References

Kyle Bagwell. *The Economic Analysis of Advertising*, volume 3. Elsevier, North Holland, 2007.

Dirk Bergemann, Benjamin Brooks, and Stephen Morris. The limits of price discrimination. *American Economic Review*, 105(3):921–957, 2015.

Dirk Bergemann, Benjamin Brooks, and Stephen Morris. Informationally robust optimal auction design. Technical report, Working Paper, Yale University, University of Chicago and Princeton University, 2016.

Dirk Bergemann, Benjamin Brooks, and Stephen Morris. First-price auctions with general information structures: Implications for bidding and revenue. *Econometrica*, 85(1):107–143, 2017.

Dirk Bergemann and Karl Schlag. Robust monopoly pricing. *Journal of Economic Theory*, 146:2527–2543, 2011.

Andrew Biehl. Durable-goods monopoly with stochastic values. *Rand Journal of Economics*, 32(3):565–577, 2001.

Simon Board. Durable-goods monoply with varying demand. *Review of Economic Studies*, 75(2):391–413, 2008.

Subir Bose, Gerhard Orosel, Marco Ottaviani, and Lise Vesterlund. Dynamic monopoly pricing and herding. *RAND Journal of Economics*, 37(4):910–928, 2006.

Subir Bose, Gerhard Orosel, Marco Ottaviani, and Lise Vesterlund. Monopoly pricing in the binary herding model. *Economic Theory*, 37: 203–241, 2008.

Benjamin Brooks and Songzi Du. Optimal auction design with common values: An informationally-robust approach. Working Paper, University of Chicago and Simon Fraser University, 2018.

Jeremy Bulow. Durable-goods monopolists. *Journal of Political Economy*, 90(2):314–332, 1982.

Rene Caldentey, Ying Liu, and Ilan Lobel. Intertemporal pricing under minimax regret. *Operations Research*, 2016.

Vinicius Carrasco, Vitor Farinha Luz, Nenad Kos, Matthias Messner, Paulo Monteiro, and Humberto Moreira. Optimal selling mechanisms under moment conditions. Technical report, Working Paper, PUC-Rio, 2017.

Gabriel Carroll. Robustness and linear contracts. *American Economic Review*, 105(2):536–563, 2015.

Gabriel Carroll. Robustness and separation in multidimensional screening. *Econometrica*, 85(2):453–488, 2017.

Sylvain Chassang. Calibrated incentive contracts. *Econometrica*, 81(5):1935–1971, 2013.

Yiwei Chen and Vivek Farias. Robust dynamic pricing with strategic customers. Technical report, Working Paper, Renmin University and MIT, 2016.

Kim-Sau Chung and Jeffrey Ely. Foundations of dominant-strategy mechanisms. *Review of Economic Studies*, 74:447–476, 2007.

John Conlisk, Eitan Gerstner, and Joel Sobel. Cyclic pricing by a durable goods monopolist. *Quarterly Journal of Economics*, 99(3):489–505, 1984.

Rahul Deb. Intertemporal price discrimination with stochastic values. Technical report, Working Paper, University of Toronto, 2014.

Songzi Du. Robust mechanisms under common valuation. *Econometrica*, Forthcoming, 2018.

Jeffrey Ely. Beeps. *American Economic Review*, 107(1):31–53, 2017.

Jeffrey Ely, Alexander Frankel, and Emir Kamenica. Suspense and surprise. *Journal of Political Economy*, 123(1):215–260, 2015.

Larry Epstein and Martin Schneider. Learning under ambiguity. *Review of Economic Studies*, 74(4):1275–1303, 2007.

Alexander Frankel. Aligned delegation. *American Economic Review*, 104(1):66–83, 2014.

Daniel F. Garrett. Intertemporal price discrimination: Dynamic arrivals and changing values. *American Economic Review*, 106(11):3275–3299, 2016.

Matthew Gentzkow and Emir Kamenica. A Rothschild-Stiglitz approach to Bayesian persuasion. *American Economic Review: Papers and Proceedings*, 106(5):597–601, 2016.

Benjamin Handel and Kanishka Misra. Robust new product pricing. *Marketing Science*, 34(6):864–881, 2015.

Justin P. Johnson and David P. Myatt. On the simple economics of advertising, marketing, and product design. *American Economic Review*, 96(3):756–784, 2006.

Anton Kolotilin. Experimental design to persuade. *Games and Economic Behavior*, 90:215–226, 2015.

Anton Kolotilin, Ming Li, Tymofiy Mylovanov, and Andriy Zapechelnyuk. Persuasion of a privately informed receiver. *Econometrica*, 2017.

Michael Landsberger and Isaac Meilijson. Intertemporal price discrimination and sales strategy under incomplete information. *Rand Journal of Economics*, 1985.

Ying Liu. Robust pricing in discrete time. Technical report, Working Paper, NYU Stern, 2016.

Alessandro Pavan, Ilya Segal, and Juuso Toikka. Dynamic mechanism design: A Myersonian approach. *Econometrica*, 82(2):601–653, 2014.

Antonio Penta. Robust dynamic implementation. *Journal of Economic Theory*, 160:280–316, 2015.

John Riley and Richard Zeckhauser. Optimal selling strategies: When to haggle, when to hold firm. *Quarterly Journal of Economics*, 98(2):267–189, 1983.

Anne-Katrin Roesler and Balasz Szentes. Buyer-optimal learning and monopoly pricing. *American Economic Review*, 2017.

Michael Rothschild and Joseph Stiglitz. Increasing risk: I. a definition. *Journal of Economic Theory*, 2(3):225–243, 1970.

Joel Sobel. Durable goods monopoly with entry of new consumers. *Econometrica*, 59(5):1455–1485, 1991.

Nancy Stokey. Intertemporal price discrimination. *Quarterly Journal of Economics*, 93(3):355–371, 1979.

Takuro Yamashita. Implementation in weakly undominated strategies: Optimality of second-price auction and posted-price mechanism. *Review of Economic Studies*, 82:1223–1246, 2015.