Informational Robustness in Intertemporal Pricing

Jonathan Libgober and Xiaosheng Mu

Department of Economics, Harvard University

May 31, 2017

Abstract. We study intertemporal price discrimination in a setting with endogenous information. A seller commits to a pricing strategy, and a buyer observes signals of her value according to some information structure. The seller does not know the information structure and thus chooses prices to maximize the worst-case profit. We show that the seller cannot do better than the one-period profit, for any number of periods and any discount factor. A deterministic constant price path delivers this optimal profit when there are arriving buyers. This paper extends the literature on informational robustness in mechanism design to a dynamic environment.

Suppose a monopolist has developed a completely new durable product and is deciding how to set prices to maximize profit. Consulting the literature on intertemporal pricing, the monopolist may at first think that if the pool of potential consumers does not change over time, profit would be maximized by charging the optimal one-period price in each period.[1] However, if the product is completely new, then the monopolist should consider the possibility that consumers will learn something about their value after pricing decisions have been made. 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. On top of this, how much journalists or product reviewers

Contact. jlibgober@fas.harvard.edu, xiaoshengmu@fas.harvard.edu. We are particularly indebted to Drew Fudenberg for guidance and encouragement. We also thank Gabriel Carroll, Ed Glaeser, Ben Golub, Johannes Hörner, Michihiro Kandori, Scott Kominers, Eric Maskin, Tomasz Strzalecki, Juuso Toikka, and our seminar audience at Harvard for comments. Errors are ours.

[1] See Stokey (1979), Bulow (1982), 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.

write about the product may depend on what prices the monopolist chooses. In general, this possibility of information arrival can make the monopolist's problem quite complicated.

This paper develops an intertemporal pricing model where a buyer observes signals of her value, possibly over time. We assume that the seller does not know the information structure (or more precisely, information arrival process) that informs the buyer of her value, and that he commits to a pricing strategy as if the information structure were the worst possible given the pricing decisions. A justification for this worst-case analysis is that the seller may want to guarantee a good outcome, no matter what the information structure actually is.[2] For our application, another justification is that a competitor may be interested in minimizing the seller's profit (even without a directly competing product). Our framework would be appropriate if other firms were able to release information on the product in a way that is outside of the seller's control.

We show that in this setting, the seller does not gain by having multiple periods to sell the object. One explanation for this result is as follows: in each period, the adversary could release information that minimizes the profit in that period. Doing so would make the seller's problem separable across time, eliminating potential gains from decreasing prices. This intuition is however incomplete, because the worst-case information structures for different periods need not be consistent, in the sense that past information may prevent the adversary from minimizing profits in the future. A key step of our argument is to show that the worst-case information structure in any period takes a partitional form, so that the seller's profit can be minimized period by period.

While selling only once achieves the optimal profit with 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 without endogenous information, 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 to the buyer 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 delay her purchase, it must also promise a greater probability of purchase to satisfy the buyer's incentives. Surprisingly, from the seller's perspective, the cost of 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, and it is in fact strictly optimal with arriving buyers.

Information arrival has long been recognized as a significant feature of many markets, and it is important to understand how it influences pricing. The difficulty, however, is that arbitrary information arrival processes could render the analysis intractable. In dynamic pricing models,

[2] 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).

it is necessary to characterize the buyer's purchasing decision, which is complicated by the interaction between prices and information. We are not aware of any papers that provide such a characterization for an *arbitrary* information arrival process. It turns out that when the seller is concerned about the worst case, the information structure that arises takes a simple, partitional form. This feature of our model enables us to describe the buyer's behavior and in turn solve for the seller's optimal pricing strategy. We hope the tools we develop will lead to simplified analysis of information dynamics in other economically meaningful settings.

In what follows, we first review the literature, and then proceed to set up the model and discuss our assumptions. In Section 3 we consider the one-period benchmark, and in Section 4 we show that a longer selling horizon does not help the seller. Section 5 shows the optimality of constant prices with arriving buyers. We discuss alternative informational and timing assumptions in Sections 6 and 7. Section 8 concludes. All omitted proofs and additional results can be found in the Appendix.

1. LITERATURE REVIEW

This paper is part of a large 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 (2017). 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.[3] Taken together, these papers characterize the seller's maxmin pricing policy and nature's minmax information structure in the *static* zero-sum game between them.[4] 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 2.1 for further discussion). Moreover, our paper is primarily concerned with dynamics, which is absent from Roesler and Szentes (2017) and Du (2017).

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

[3]Du (2017) extends the analysis to a one-period, many-buyer common value auction environment. He constructs a class of mechanisms that extracts full surplus when the number of buyers grows to infinity, despite the presence of informational uncertainty. However, which mechanism achieves the maxmin profit remains an open question for finitely many buyers. This is solved in the special case of two buyers and two value types by Bergemann, Brooks and Morris (2016).

[4]Roesler and Szentes (2017) actually motivate their model as one where the buyer chooses the information structure; they show that this solution also minimizes the seller's profit.

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.[5] 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 (in a one-period model) that a deterministic price is optimal when the seller only knows the true value distribution to be in some neighborhood of distributions.[6] 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. In our setting, information arrival places restrictions on *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.

Absent robustness concerns, several intertemporal pricing papers 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) and Garrett (2016) allow for stochastically changing values, but in these papers the evolution of values violates the martingale condition for expectations.[7] As stated above, the maxmin objective leads us to the study of simple and intuitive information structures, making the buyer's problem tractable. While we believe that a Bayesian version of our problem is worth studying, we are not aware of how to determine a buyer's optimal purchasing behavior under an arbitrary information arrival process.

Finally, the implication of dynamic information arrival has been considered in the related literature of information design. The connection arises because this literature seeks to describe how a receiver's (buyer) behavior varies depending on how a sender (nature) chooses the information

[5]These authors do not allow nature to condition on the realized price, so their paper focuses on the alternative timing that we discuss in Section 7.

[6]Their result applies to maxmin profit as in our model. The authors also show that if the seller's objective is instead to minimize regret, then random prices do better.

[7]Deb (2014) assumes the value is independently redrawn upon Poisson shocks. For Garrett (2016), the value follows a two-type Markov-switching process.

structure, see Ely, Frankel and Kamenica (2015) and Ely (2017). Since we are ultimately concerned with pricing strategies by the seller, we cannot directly borrow the techniques from these papers. However, 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.

2. MODEL

A seller (he) sells a durable good to a buyer (she)[8] at time $t = 1, 2, \ldots, T$, where $T \leq \infty$. Both the seller and the buyer discount the future at rate δ. The product is costless for the seller to produce,[9] while the buyer has unit demand and obtains discounted lifetime utility from purchasing the object equal to v. The *value* v has distribution F supported on $\mathbb{R}_+$, with $0 < \mathbb{E}[v] < \infty$. We let $\underline{v}$ denote the minimum value in the support of F. The distribution F is fixed and common knowledge, and the buyer's value for the object does not change over time.

However, the buyer does not directly know v; instead, she observes signals which give information about v. An information structure is a function which maps true values into realizations of signals. To be precise, a *dynamic information structure* $\mathcal{I}$ is a sequence of signal sets S_t and probability distributions $I_t : R_+ \times S^{t-1} \times P^t \to \Delta(S_t)$, for $1 \leq t \leq T$.[10] The interpretation is that the buyer observes signal realization s_t at time t, whose distribution depends on her true value $v \in R_+$, the history of previous signal realizations $s^{t-1} = (s_1, s_2, \ldots, s_{t-1}) \in S^{t-1}$, as well as the history of previous *and current* period prices $p^t = (p_1, p_2, \ldots, p_t) \in P^t$.

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. Then, nature chooses a dynamic information structure. In each period $t \geq 1$, the price in that period p_t is realized according to $\sigma(p_t \mid p^{t-1})$. The buyer with true value v observes the signal s_t with probability $I_t(s_t \mid v, s^{t-1}, p^t)$ and decides whether or not to purchase the product.

Given the pricing strategy σ and the information structure $\mathcal{I}$, the buyer faces an optimal stopping problem. Specifically, she chooses a stopping time τ^* adapted to the joint process of

[8] Our analysis is unchanged in the case of many identical buyers who do not know their true values, as the worst case would simply involve each buyer being given the same information structure. The case with arriving buyers is considered in Section 5.

[9] 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.

[10] To avoid measurability issues, we assume each signal set S_t is at most countably infinite. All information structures in our analysis have this property.

prices and signals, so as to maximize the expected discounted value less price:

$$\tau^* \in \operatorname*{argmax}_{\tau} \mathbb{E}\left[\delta^{\tau-1}(\mathbb{E}[v|s^\tau, p^\tau] - p_\tau)\right].$$

The inner expectation $\mathbb{E}[v|s^\tau, p^\tau]$ 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 evolution of prices and signals. We note that the stopping time τ is allowed to take any positive integer value $\leq T$, or $\tau = \infty$ to mean the buyer never buys.

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

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

Note that when the 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.[11]

2.1. 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, the buyer knows the information structure, and she 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 rely upon some product review website and hence know very well how to interpret the reviews. The seller, on the other hand, knows that there are many possible ways the buyer can learn about her value, and he wants to do well against all these possibilities. In Section 6, we will show that our results extend even if the seller knows the buyer begins with extra prior information (say, through advertising). 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 classic

[11] 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.

setting where the buyer knows her value can be seen as an extreme case of our extended model, where the buyer has a more informative prior and may receive additional information over time.

Third, our key timing assumption is that the buyer's information structure in each period is determined *after* the price for that period has been realized. One could also consider an alternative model where, in each period, nature first chooses an information structure that depends only on past prices, and subsequently the current price is realized. We provide an analysis of this alternative timing in Section 7. As stated above, the one-period optimal seller strategy in this model follows from Roesler and Szentes (2017) and Du (2017). In the dynamic setting, we are able to generalize the robust selling mechanism of Du (2017) and prove its optimality when each buyer only receives information upon arrival. However, we view our timing assumption as more appropriate for the applications mentioned in the Introduction, where the seller is restricted to posted prices. For example, the seller may update his price at the beginning of every quarter, and any buyer can purchase at that price in the next three months. What information (e.g. product reviews or competitors' advertisements) buyers will receive in such situations may well depend on the realized price.

Finally, we assume that the seller commits to a pricing strategy that the buyer observes, and it is independent of nature's choice of the information structure. In particular, the seller must have some method of committing to (and communicating to the buyer) a particular randomization in the future; although, as it turns out, randomization does not help the seller under our timing assumption. Studying robust intertemporal pricing with limited commitment is left to future work.

3. ONE-PERIOD BENCHMARK

We start with the case where the seller only has one period to sell the object. To solve this problem, we define a transformed distribution of F. For expositional simplicity, the following definition assumes F is continuous. All of our results in this paper extend to the discrete case, 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.*

The distribution G is useful because for any (realized) price p, nature can only ensure that the object remains unsold with probability $G(p)$. This holds since the worst-case information structure has the property that a buyer who does not buy has expected value exactly p. Otherwise nature could find an information structure that makes a non-purchasing buyer slightly more optimistic

about her value and decreases the probability of sale. This observation allows us to show that the worst-case information structure involves telling the buyer whether her value is above or below $F^{-1}(G(p))$, making $1 - G(p)$ the probability of sale.

This discussion gives 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:*

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

For future reference, we call p^* the one-period maxmin price and similarly $\Pi^* = \max_p p(1 - G(p))$ 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 transformed distribution G takes the place of F, which will be useful for the analysis in later sections. The following example illustrates:

Example 1. *Let $v \sim$ Uniform[0,1], so that $G(p) = 2p$. 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 $v > 1/2$.*

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

4. MULTIPLE PERIODS

In this section we present our main result that having multiple periods to sell does not improve the seller's profit when nature can release information dynamically.[12] Since the seller can always sell exclusively in the first period, the one-period profit Π^* forms a lower bound for the seller's maxmin profit.

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

[12] We highlight 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.

decomposes into a convex combination of one-period profits. Our proof takes advantage of the partitional form of worst-case information structures to show that nature can minimize the seller's profit period by period.

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 profit no more than Π^*. Hence the seller's maxmin profit against all dynamic information structures is Π^*, irrespective of the time horizon T and the discount factor δ.*

We will present the proof of this proposition under the 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 $F(w_{t_{j-1}}) - F(w_{t_j})$.

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 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.

We describe the buyer's optimal stopping behavior 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 proof of this lemma can be found in Appendix A, where we actually prove a more 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. 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 (assuming $T = \infty$):

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

where the second line is by Abel summation,[13] the third line is by w_{t_j}'s indifference between buying in period t_j or t_{j+1}, and the last inequality uses $w_{t_j}(1 - G(w_{t_j})) \leq \Pi^*, \forall j$. For finite horizon T, the proof proceeds along the same lines except for a minor modification to Abel summation. ■

While most of the intuition for the general result is captured by the above argument, random prices introduce a technical difficulty in applying the sorting argument directly. Specifically, since the threshold values w_{t_j} depend on both the realized price and the distribution of future prices, they are in general random variables. More problematically, these thresholds may be non-monotonic if they are to be defined using the buyer's indifference condition. If 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 above proof will fail.

To recover the proof, the trick we use (the details of which are in Appendix A) is to consider modified thresholds that are *forced* to be decreasing. That is, we define v_t to 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 $w_t = \min\{v_1, v_2, \ldots, v_t\}$. Using this modified definition for w_t, we can consider the same information structure as in the above proof and show that Proposition 2 continues to hold for random prices.

[13] Abel summation says that $\sum_{j\geq 1} a_j b_j = \sum_{j\geq 1} \Big((a_j - a_{j+1}) \sum_{i=1}^{j} b_i\Big)$ 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 $b_j = G(w_{t_{j-1}}) - G(w_{t_j})$.

5. ARRIVING BUYERS

One may wonder if the conclusions we have derived would continue to hold if buyers were to arrive over time, as in Conlisk, Gerstner and Sobel (1984), Sobel (1991), Board (2008) and Garrett (2016). In contrast to the single buyer case, selling only once is no longer optimal as the monopolist may want to capture buyers who only arrive in later periods. On the other hand, low prices in the future could allow nature to choose an information structure that delays purchase, which may be costly for the seller.

To comment on this possibility, we modify the model by assuming that in each period t, a new buyer arrives and decides when to buy the object. To be precise, our timing is as follows:

- At time 0, the seller chooses a pricing strategy $\sigma \in \Delta(p^T)$.
- At time $t \in \{1, \ldots, T\}$, a buyer arrives with value $v^{(t)}$ drawn from F, independently from previous buyers.
- Once a buyer arrives at time t, nature chooses a dynamic information structure $\mathcal{I}^{(t)}$ according to which this buyer learns her value.[14]
- Given the pricing strategy σ and the information structure $\mathcal{I}^{(t)}$, the buyer arriving at time t chooses a stopping time $\tau^{(t)}$ to purchase the object.

By Proposition 2, the seller's discounted profit from the buyer arriving at time t is bounded above by $\delta^{t-1} \cdot \Pi^*$. Thus, an upper bound for overall profit is $\frac{1-\delta^T}{1-\delta}\Pi^*$. If the seller were able to set personalized prices, this upper bound could be achieved by selling only once to each arriving buyer. Surprisingly, 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 (if she were to buy at all), due to impatience. However, the promise of future information may induce delay. In the following lemma, we show that the seller can eliminate the potential damage of delayed purchase by committing to never lower the price.

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

[14]Because buyers have independent values, whether these information structures are private to each individual buyer or public to all buyers does not affect the analysis. It may be of interest to study the case where values are correlated, and nature/adversary is restricted to releasing public information. We leave this extension for future work.

We present the intuition here and leave the formal proof to Appendix A. Let us fix a non-decreasing price path. For any dynamic information structure nature can choose, we consider an alternative information structure that simply informs the buyer of her stopping time, in the first period. This replacement is in the spirit of the revelation principle; however, it differs due to the fact that we push nature's recommendation to time 1. The proof shows that for non-decreasing prices, we can find a replacement such that the buyer still follows nature's recommendation of whether or not to buy. This replacement has the property that the seller's profit is decreased. Since the seller receives at least Π^* under any information structure that releases information only in the first period, we obtain the lemma.

Armed with this lemma, we can show the following:

Proposition 3. *In the multi-period model with arriving buyers, the seller can guarantee* $\frac{1-\delta^T}{1-\delta}\Pi^*$ *with a constant price path charging* p^* *in every period. This deterministic pricing strategy is optimal, and it is uniquely optimal whenever the one-period maxmin price* p^* *is unique.*

6. INITIAL BUYER INFORMATION

Our model so far assumes that the seller has no knowledge over the information the buyer receives. In practice, however, the seller may know that the buyer has access to at least some information. For example, he may conduct an advertising campaign, and understand its informational impact very well (Johnson and Myatt (2006)). While it may be impossible or difficult for such an advertising campaign to remove all uncertainty, the seller may nevertheless know that the buyer has access to some baseline information.[15] In this section we show that this possibility does not change our conclusions.

We modify the model in Section 2 by assuming that in addition to having the prior belief F, the buyer observes some signal $s_0 \in S_0$ at time 0. The signal set S_0 as well as the conditional probabilities of s_0 given v are common knowledge between the buyer and the seller, and we denote this initial information structure by $\mathcal{H}$. We allow nature to provide information conditional on s_0 but keep all other aspects of the model identical. Equivalently, the seller seeks to be robust against all dynamic information structures in which buyer learns $\mathcal{H}$ and possibly more information in the first period.

A signal s_0 induces a posterior belief on the buyer's value, which we denote by the distribution F_{s_0}. Define G_{s_0} to be the transformed distribution of F_{s_0}, following Definition 1. The same analysis

[15]Note that if the seller has complete control over what information he provides, it would be impossible to do better than the full information outcome because nature could always reveal the value.

as in Section 3 yields the following result:

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:*

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

We denote the maxmin profit in this case by $\Pi^*_{\mathcal{H}}$.

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 $\mathbb{E}_{s_0}[G_{s_0}(p)] = F(p)$. In contrast, if $\mathcal{H}$ is completely uninformative, we return to Equation (1).

For the multi-period problem, our previous proof also carries over and shows that the seller does not benefit from a longer selling horizon.

Proposition 2'. *In the multi-period model where the buyer observes initial information structure $\mathcal{H}$, the seller's maxmin profit against all dynamic information structures is $\Pi^*_{\mathcal{H}}$, irrespective of the time horizon T and the discount factor δ.*

The proofs of these results are direct adaptations of those for the model without an initial information structure. Thus we omit them from the Appendix.

7. ALTERNATIVE TIMING

We have assumed that in each period, nature can release information that depends on realized prices in previous periods *as well as* in the current period. Here we consider an alternative timing of the model, where nature only conditions on past prices. Formally, throughout this section we re-define a *dynamic information structure* to be a sequence of signal sets $(S_t)_{t=1}^T$ and probability distributions $I_t : R_+ \times S^{t-1} \times P^{t-1} \to \Delta(S_t)$. The crucial distinction from our main model is that the signal s_t depends on previous prices p^{t-1} but not on the current price p_t. The seller chooses a pricing strategy that achieves maxmin profit against such information structures and corresponding optimal stopping times of the buyer.

With a single period, this model reduces to one studied in Roesler and Szentes (2017) and Du (2017): the seller and nature play a zero-sum game in which the seller chooses a distribution of prices (equivalently, a mechanism), while nature chooses an information structure. To make the connection 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 $\tilde{F}$.[16] 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 for some W, B, such that W is smallest possible subject to F being a mean-preserving spread of F_W^B.

It follows that the seller's one-period profit is at most the smallest W defined above, which we denote by Π_{RSD}. Conversely, Du (2017) constructs a particular mechanism the seller can use to guarantee profit at least Π_{RSD} under any information structure nature chooses. In Appendix B, we represent Du's "exponential mechanism" as an equivalent random price mechanism. The results of Roesler-Szentes and Du together imply that Π_{RSD} is the one-period maxmin profit. We note that $\Pi_{RSD} \geq \Pi^*$ in general, and in Appendix B we characterize when the inequality is strict.

With multiple periods, 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 at most $W = \Pi_{RSD}$. To summarize, we have shown:[17]

Proposition 4. *Suppose there is a single buyer. For any time horizon T and any discount factor δ, the seller's maxmin profit when nature cannot condition on the current period price is given by Π_{RSD}.*

Lastly we consider the case of arriving buyers. The setup is identical to Section 5, except that nature is now restricted to condition on past prices only. Whether or not the seller can obtain

[16] This equivalence is separately observed by Gentzkow and Kamenica (2016) in the context of Bayesian persuasion. These authors attribute the result to Rothschild and Stiglitz (1970).

[17] It is worth comparing Proposition 4 to Proposition 2. Both results show that regardless of the timing of nature's moves, a longer selling horizon does not help the seller. Here, this conclusion follows from the duality between Roesler-Szentes and Du—as the above proof shows, Proposition 4 continues to hold even if nature only provides information in the first period. In our main model however, nature had to counter every pricing strategy with a *dynamic* information structure.

profit Π_{RSD} from each arriving buyer turns out to depend on the dynamics of information arrival. Specifically, this profit benchmark can be achieved if nature only releases information to a buyer when she arrives.

Proposition 5. *Suppose there are arriving buyers, and suppose each buyer only receives information once upon arrival (before the price realizes in that period). For any time horizon T and any discount factor δ, the seller has a pricing strategy that ensures profit at least Π_{RSD} from each buyer. Thus the seller's maxmin profit is $\frac{1-\delta^T}{1-\delta}\Pi_{RSD}$.*

One can interpret Proposition 5 as the following dynamic extension of Roesler-Szentes' model: each arriving buyer enters the market with potentially more information than the prior F, but there is no learning over time. Our result generalizes Du's static mechanism to such dynamic settings, showing a pricing strategy exists that is robust not only to buyers' information, *but also to their arrival times.* The proof is based on a key lemma (Lemma 4 in Appendix B) relating the outcome under a static price distribution to that under a dynamic price distribution. This outcome-equivalence property enables us to construct a dynamic pricing strategy that replicates Du's mechanism for each arriving buyer, achieving Π_{RSD} as profit guarantee.

Nevertheless, our construction for Proposition 5 is not robust to *dynamic* information arrival. The following result shows that in general, the seller cannot obtain Π_{RSD} from each arriving buyer who may learn over time.

Proposition 6. *Consider a model with two periods and one buyer arriving in each period. Suppose nature can provide information dynamically (to the first buyer). Assume that $\Pi_{RSD} > \Pi^*$ and that Du's mechanism is uniquely maxmin optimal in the one-period problem. Then the seller's maxmin profit in this two-period model with arriving buyers is strictly below $(1+\delta)\Pi_{RSD}$ for any $\delta \in (0,1)$.*

While the proof of this proposition is fairly complicated, the information structure chosen by nature is simple. 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. We show that nature can induce delayed purchase from the first buyer and further damage profit by promising future information. Specifically, nature can reveal the value perfectly, in the second period, to any buyer who *would have* purchased in the first period without any additional information. The key technical step of the proof shows that delay always hurts the seller, and it occurs with strictly positive probability.[18]

[18] We are only able to show that for this specific information structure, total profit is strictly below $(1+\delta)\Pi_{RSD}$. Since this 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 model considered here.

This last statement relies on our assumption that $\Pi_{RSD} > \Pi^*$: as we showed in Lemma 2 for the reverse timing, if the seller charges a *deterministic* constant price path, nature cannot hurt the seller with the promise of future information. Proposition 6 can thus be interpreted as saying that whenever randomization is required, the one-period profit benchmark Π_{RSD} is unattainable with arriving buyers and dynamic learning. In this sense we view Π^* as a more cautious benchmark even under the timing assumption discussed here.

On the other hand, we have stated Proposition 6 with 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 this assumption holds for *generic* F.

8. CONCLUSION

In this paper, we have studied optimal monopoly pricing in relation to dynamic information arrival, utilizing a maxmin robustness approach to provide a sharp answer on how they interact. We show that the monopolist's optimal profit is what he would obtain with only a single period to sell to each buyer. Furthermore, a constant price path delivers this optimal profit, even when buyers arrive over time. The inability to condition on a buyer's arrival time therefore imposes no cost on the seller (in our main model). While these results have long been known in cases where buyers know their values, the profit and prices in our model are typically lower.

The main lesson of this paper—namely, the optimality of constant price paths and their corresponding profit level—is one that we hope will continue to be scrutinized in other contexts and under other modeling assumptions. The case of limited seller commitment seems compelling, though there are technical difficulties associated with formalizing (seller) learning under ambiguity (see Epstein and Schneider (2007)). One could also ask similar questions in more general dynamic mechanism design settings, where the agent's problem may not be represented by the choice of a stopping time.

This paper contributes to a growing literature which employs the maxmin approach in analyzing the optimal design of mechanisms. In our setting, the maxmin approach allows us to focus on particular information structures—the partitional ones—with clear economic interpretations. We therefore avoid the difficulty in working with the entire space of information structures and stopping times. While it is certainly worthwhile to analyze the Bayesian model, doing so would first require a similarly tractable restriction as the one we have provided here.

A. PROOFS FOR THE MAIN MODEL

We first define the transformed 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 $\beta > \mathbb{E}[v]$.[19]

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 that the buyer either "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 $v \leq \alpha$. From this, it is 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.

[19] 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

$$v - p_t \geq \max_{\tau \geq t+1} \mathbb{E}[\delta^{\tau - t} \cdot (v - p_\tau)] \tag{5}$$

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 $v \geq v_t$.[20] Thus, v_t is defined by the equation

$$v_t - p_t = \max_{\tau \geq t+1} \mathbb{E}[\delta^{\tau - t} \cdot (v_t - p_\tau)] \tag{6}$$

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

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

$$w_t = \min\{v_1, v_2, \ldots, v_t\} = \min\{w_{t-1}, v_t\}. \tag{7}$$

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 (7). Given this signal, she knows that she will receive no more information in the future (because w_t decreases over time). She also knows that her value is above the $G(w_t)$-percentile, which is greater than $w_t = v_t$, the average value below that percentile. Thus from the definition of v_t, the buyer optimally buys in period t.

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

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 is Blackwell sufficient for all previous signals. By definition of G, the buyer's expected value is $w_t \leq v_t$. Thus even without additional information in the future, this buyer prefers to delay her purchase. The promise of future information does not change the result. ■

Step 3: Profit decomposition. By this lemma, the buyer with true value in the percentile range $(G(w_{t-1}), G(w_t)]$ buys 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 3. *Suppose $w_t = v_t \leq w_{t-1}$ in some period t. Then*

$$p_t = \mathbb{E}\left[\sum_{s=t}^{T-1}(1-\delta)\delta^{s-t}w_s + \delta^{T-t}w_T \mid p^t\right] \tag{8}$$

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

Using Lemma 3, we can rewrite the profit as

$$\begin{aligned}
\Pi &= \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} \cdot (G(w_{t-1}) - G(w_t)) \cdot \mathbb{E}\left[\sum_{s=t}^{T-1}(1-\delta)\delta^{s-t}w_s + \delta^{T-t}w_T \mid p^t\right]\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} \cdot (G(w_{t-1}) - G(w_t)) \cdot \left(\sum_{s=t}^{T-1}(1-\delta)\delta^{s-t}w_s + \delta^{T-t}w_T\right)\right] \\
&= \mathbb{E}\left[\sum_{s=1}^{T-1}(1-\delta)\delta^{s-1}w_s(1-G(w_s)) + \delta^{T-1}w_T(1-G(w_T))\right] \\
&\leq \Pi^*.
\end{aligned} \tag{9}$$

The second line is by the law of iterated expectations, because 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. Hence it only remains to prove Lemma 3.

Proof of Lemma 3. We assume that T is finite. The infinite-horizon result follows from an approximation by finite horizons and the Monotone Convergence Theorem, whose details we omit. We

prove by induction on $T-t$, where the base case $t=T$ follows from $w_T = v_T = p_T$. For $t<T$, from (6) 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

$$p_t = \mathbb{E}[(1-\delta^{\tau-t})v_t + \delta^{\tau-t}p_\tau]. \tag{10}$$

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 (7); 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 (10) 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.[21] To prove this, note that since stopping at time τ is weakly better than stopping at time 1, we have

$$\mathbb{E}[v \mid s_1] - p_1 \leq \mathbb{E}^{s_2,\cdots,s_T}\left[\delta^{\tau-1}(\mathbb{E}[v \mid s_1, s_2, \cdots, s_\tau] - p_\tau)\right]. \tag{11}$$

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 (11) to the following.

$$\mathbb{E}[v \mid s_1] \leq \mathbb{E}^{s_2,\cdots,s_T}\left[\delta^{\tau-1}\mathbb{E}[v \mid s_1, s_2, \cdots, s_\tau] + (1-\delta^{\tau-1})p_1\right]. \tag{12}$$

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:

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

The denominator $\mathbb{E}^{s_2,\cdots,s_T}\left[1-\delta^{\tau-1}\right]$ 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 (13) 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 $s_1, \ldots, s_T$.

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

With these, (13) states that

$$p_1 \geq \frac{\mathbb{E}^{s_2,\cdots,s_T}\left[\mathbb{P}(\underline{s} \mid s_1, s_2, \ldots, s_T) \cdot \mathbb{E}[v \mid s_1, s_2, \cdots, s_T, \underline{s}]\right]}{\mathbb{E}^{s_2,\cdots,s_T}\left[\mathbb{P}(\underline{s} \mid s_1, s_2, \ldots, s_T)\right]} = \mathbb{E}[v|s_1, \underline{s}] \tag{14}$$

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 Proposition 3

By the previous lemma, 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, then from (9) we see that the seller's profit from the first buyer equals Π^* only if $w_s = p^*$ almost surely. This together with Lemma 3 implies $p_1 = p^*$ almost surely. Analogous argument for later buyer shows that the seller must always charge p^* to achieve the maxmin profit. Hence the proposition.

A.5. 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 7. *For any continuous distribution F, let $\hat{p}$ be an optimal monopoly price under known values:*

$$\hat{p} \in \operatorname*{argmax}_p p(1 - F(p)). \tag{15}$$

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

Proof of Proposition 7. 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,

$$pG(p) = \int_0^{F^{-1}(G(p))} v dF(v), \forall p \in [\underline{v}, \mathbb{E}[v]]. \tag{16}$$

Differentiating both sides with respect to p, we obtain

$$G(p) + pG'(p) = \frac{\partial}{\partial p}(F^{-1}(G(p))) \cdot F^{-1}(G(p)) \cdot F'(F^{-1}(G(p))) = G'(p) \cdot F^{-1}(G(p)). \tag{17}$$

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

$$G'(p) = \frac{G(p)}{F^{-1}(G(p)) - p}. \tag{18}$$

Thus,

$$G(p) + pG'(p) = \frac{G(p) \cdot F^{-1}(G(p))}{F^{-1}(G(p)) - p}. \tag{19}$$

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

$$\frac{p}{1 - G(p)} \cdot \left(1 - F\left(\frac{p}{1 - G(p)}\right)\right) < p. \tag{20}$$

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. ■

B. PROOFS FOR THE ALTERNATIVE TIMING MODEL

B.1. Comparison Between Π^* and Π_{RSD}

In what follows we focus on the alternative model described in Section 7, where nature cannot condition on the current period price. We show that the relevant profit benchmark Π_{RSD} is in general higher than Π^*, and the difference may be significant:

Proposition 8. *$\Pi_{RSD} \geq \Pi^*$ 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. 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.[22] ■

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

B.2. Proof of Proposition 5

Throughout, we represent the robust selling mechanism in Du (2017) by a random price, with c.d.f. $D(x)$; the details of this distribution can be found later in (31), but they are not relevant for this proof. Because nature can provide each arriving buyer with the Roesler-Szentes information structure (4), the seller at most obtains Π_{RSD} from each buyer. To complete the proof, we will construct a dynamic pricing strategy that yields Π_{RSD} from each buyer.

The following lemma proves the outcome-equivalence between static and dynamic pricing strategies, and it may be of independent interest:

Lemma 4. *Fix any continuous distribution function D, 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 outcome as the static prices for every type of each arriving buyer.

We state the lemma for continuous distributions so that the buyer's optimal stopping time is almost surely unique. From Du (2017), Du's distribution D is continuous except when it is a point-mass on W. In the latter case $\Pi_{RSD} = \Pi^*$, and Proposition 5 follows from Proposition 3.

Lemma 4 is useful for our problem because it implies, via the Revenue Equivalence Theorem, that 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 distributed according to 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 proposition. Since Du's static mechanism guarantees profit Π_{RSD} from every buyer, the proposition will follow once we prove the lemma.

Proof of Lemma 4. 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,

$$v_1 - p_1 = \delta \cdot \mathbb{E}^{p_2 \sim D} \left[\max\{v_1 - p_2, 0\} \right]. \tag{21}$$

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 purchasing 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 $(1-\delta)(1-H(x)) + \delta(1-H(x))(1-D(x))$.[23] This has to be equal to $1 - D(x)$, which implies

$$1 - H(x) = \frac{1 - D(x)}{1 - \delta D(x)}. \tag{22}$$

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 (21)). We have proved the lemma for $T = 2$.

Before proceeding, we remark that (22) implies the distribution H has the same support as D. However, (21) 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 case we have discussed). Intuitively, charging this maximum price in the first period leads to delayed purchase by buyers with high values, which is costly for the seller. On the other hand, the minimum price p_1 is indeed equal to the minimum of the support of D, which we denote by W; when D is Du's distribution, this is the same W as in the Roesler-Szentes information structure (4).

Step 2: Extension to finite T. We conjecture a pricing strategy σ that is independent across periods: $d\sigma(p_1, \ldots, p_T) = d\sigma_1(p_1) \times \cdots \times d\sigma_T(p_T)$, where we interpret each σ_t as a distribution. Define the cutoff values $v_1, \ldots, v_T$ as in (6). 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 purchasing probability is $\mathbb{P}[w^{(t)} \leq x]$. The random variable $w^{(t)}$ is described as follows: for $t \leq s \leq T-1$, $w^{(t)} = \min\{v_t, v_{t+1}, \ldots, v_s\}$ with probability $(1-\delta)\delta^{s-t}$; and with remaining probability δ^{T-t}, $w^{(t)} = \min\{v_t, v_{t+1}, \ldots, v_T\}$.

[23] $1 - H(x)$ is the probability that $w_1 > x$, and $(1-H(x))(1-D(x))$ is the probability that $w_2 > x$.

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 $v_1, \ldots, v_{T-1} \sim H$ (since v_t depends only on p_t, they are independent random variables).[24] We can then solve for the price distributions $\sigma_1, \ldots, \sigma_T$ by backward induction: σ_T must be D, and once the prices in period $t+1, \ldots, T$ are determined, there is a one-to-one relation between p_t and v_t by (6). 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 $\sigma_1, \sigma_2, \ldots$ such that $v_1, v_2, \cdots \sim H$. We conjecture a stationary σ_t. Recall that the cutoff v_1 is defined by

$$v_1 - p_1 = \max_{\tau \geq 2} \mathbb{E}\left[\delta^{\tau-1}(v_1 - p_\tau)\right]. \tag{23}$$

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 $v_1 - p_1$. (23) 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

$$v_1 = p_1 + \frac{\delta}{1-\delta} \cdot \mathbb{E}^{p_2}\left[\max\{p_1 - p_2, 0\}\right]. \tag{24}$$

Let $P(x)$ denote the c.d.f. of p_1 (and of p_2). When $p_1 = x$, (24) implies

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

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

$$P(x) = H\left(x + \frac{\delta}{1-\delta} \int_0^x P(z)\, dz\right). \tag{25}$$

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

$$Q(x) = x + \frac{\delta}{1-\delta} \int_0^x P(z)\, dz; \qquad U(y) = 1 + \frac{\delta}{1-\delta} H(y) = \frac{1}{1 - \delta D(y)}. \tag{26}$$

[24] 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 (22).

(25) is the differential equation

$$U(Q(x)) = Q'(x). \tag{27}$$

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

$$\frac{\partial V(Q(x))}{\partial x} = V'(Q(x)) \cdot Q'(x) = \frac{Q'(x)}{U(Q(x))} = 1. \tag{28}$$

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 (28) that $Q(x)$ is characterized by

$$V(Q(x)) = x \text{ with } V(y) = \int_0^y (1 - \delta D(z))\, dz. \tag{29}$$

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

$$P(x) = \frac{1-\delta}{\delta} \cdot (Q'(x) - 1). \tag{30}$$

Lemma 4 is proved, and so is Proposition 5. ■

B.3. Proof of Proposition 6

The proof is somewhat long, and we will present it in several steps. First, we review some properties of Du's static mechanism. Next, we focus on the pricing strategy σ^D that we constructed in the preceding proof. We construct a *dynamic* information structure (for the first buyer) that yields profit below Π_{RSD}. This proves the proposition assuming that the seller uses the strategy σ^D. Lastly, we apply continuity arguments and extend the result to any pricing strategy σ.

Step 1: Properties of Du's mechanism. For the one-period model, Du (2017) constructs a mechanism that guarantees profit Π_{RSD} regardless of the buyer's information structure. By considering the profile of interim allocation probabilities as a c.d.f., we can equivalently implement Du's mechanism as a random price with the following distribution:

$$D(x) = \begin{cases} 0 & x \in [0, W) \\ \frac{\log \frac{x}{W}}{\log \frac{S}{W}} & x \in [W, S) \\ 1 & x \in [S, 1] \end{cases} \tag{31}$$

Here $S \in (W, B]$ is characterized by[25]

$$\int_0^S F_W^B(v)\, dv = \int_0^S F(v)\, dv \tag{32}$$

where F_W^B is the Roesler-Szentes worst-case information structure (4). To explain further, Roesler and Szentes (2017) observe that the LHS in (32) must not exceed the RHS (for all S) because F is a mean-preserving spread of F_W^B. However, when W is smallest possible, this constraint must bind at some S.

The following observations will be crucial. 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 in fact 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, the Roesler-Szentes information structure has the property that a buyer with true value $v \leq S$ only receives signal $\leq S$ (i.e. her posterior expected value is at most S), while a buyer with true value $v > S$ expects her value to be greater than S.

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 (32) and the last one uses (4).

We note that in general, there could be more than one point S for which (32) holds. Thus, the maxmin optimal mechanism in one period need not be unique even if we restrict attention to the class of exponential mechanisms considered by Du (2017). But we do have the following result:

Lemma 5. *There is a unique maxmin optimal mechanism in the one-period simultaneous-move model if and only if (32) holds at a unique point S.*

We mention that for generic distributions F, there is a unique S that satisfies (32). However, the proof is tangential to the paper and we will leave it out. A sufficient condition is that $F(x)$ is

[25] S is strictly greater than W because otherwise D is a mass-point at W and $\Pi_{RSD} = \Pi^*$, contradicting the assumption of the proposition.

convex, for example when F is uniform.[26]

Proof of Lemma 5. "Only if" is obvious, so 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. Suppose $r(p)$ is the p.d.f. of the random price, then profit is

$$\Pi = \int_0^1 p \cdot r(p) \cdot (1 - \tilde{F}(p))\, dp. \tag{33}$$

Given $r(p)$, Nature's problem is to choose a c.d.f. $\tilde{F}$ to minimize Π, subject to $\int_0^x \tilde{F}(v)\, dv \leq \int_0^x F(v)\, dv$ for all $x \in (0, 1]$, with equality at $x = 1$ (so that $\tilde{F}$ has the same mean as F).

By Roesler and Szentes (2017), $\tilde{F} = F_W^B$ is a solution to nature's problem. For this solution, the integral inequality constraint only binds at $x = S$. Standard perturbation techniques in the calculus of variations thus imply that $\tilde{F} = F_W^B$ cannot be improved upon only if $p \cdot r(p)$ is a constant for $p \in (W, S)$.[27] Similarly, $p \cdot r(p)$ must also be a constant on the interval $p \in (S, B)$; in fact, we can show this constant is zero.[28]

Hence, $r(p)$ 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 D. ■

Step 2: The Information Structure. Consider now the model with two periods and one buyer arriving in each period. The problem for the second buyer is static, so nature can choose an information structure that yields profit at most Π_{RSD}.

We construct the following dynamic information structure $\mathcal{I}$ for the *first* buyer:

- In the first period, nature provides the Roesler-Szentes information structure. We denote the buyer's unbiased signal by $\tilde{v}$ (which is also her posterior expected value), 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 as well as the buyer's expected value $\tilde{v}$ in the first period, nature reveals the buyer's true value v if and only if $\tilde{v} \geq v_1(p_1)$.

[26] Recall that $F(S) = F_W^B(S)$. However, $F(x) - F_W^B(x) = F(x) + \frac{W}{x} - 1$ is convex, and so it has at most two roots $x_0 < x_1$. Because $F(x) > F_W^B(x)$ for $x < x_0$, S being x_0 would contradict (32). Thus $S = x_1$ is unique.

[27] Suppose to the contrary 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 $\tilde{F}$ still satisfies the feasibility constraints, and the profit Π is reduced.

[28] If this constant were $c > 0$, then on the interval $[S, B]$ nature seeks to minimize $c \cdot \int_S^B (1 - \tilde{F}(v))\, dv$ subject to the integral inequality constraint and equal means: $\int_S^1 (1 - \tilde{F}(v))\, dv = \int_S^1 (1 - F(v))\, dv$. Thus nature equivalently maximizes $\int_B^1 (1 - \tilde{F}(v))\, dv$. Choosing $\tilde{F} = F_W^B$ results in 0 and is sub-optimal.

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:

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

Note that in general, the distribution of p_2 may depend on p_1.

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 3: Buyer behavior and seller profit. To facilitate the discussion, we consider another 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. Secondly, 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 situation cannot occur because $\mathcal{I}$ provides more information than $\mathcal{I}'$ in the second period, and the buyer's incentive to wait could only be stronger.

Thus, when nature provides $\mathcal{I}$ rather than $\mathcal{I}'$, the seller's profit changes only in the last possibility above. Let us show that whenever the buyer delays her purchase from the first period to the second, the seller's profit decreases by at least $(1-\delta)W$. This is because when the buyer chooses to not buy in the first period, the discounted social surplus decreases by at least $(1-\delta)\tilde{v}$. Since the buyer's payoff cannot decrease (because she *chooses* to delay purchase), the loss must come from the seller's discounted profit.

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 at least smaller by $(1-\delta)W$ times the probability that the buyer delays purchase.*

Step 4: Proof of the proposition for σ^D. Let σ^D be the pricing strategy given by Lemma 4, which we recall is robust to information that arrives only once (for each buyer). Here we show that under the dynamic information structure $\mathcal{I}$, the seller's profit from the first buyer is strictly less than Π_{RSD}.

Recall from the proof of Lemma 4 that under σ^D, the price in the second period p_2 is drawn from Du's distribution D, independent 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 (22)).

Suppose the buyer receives unbiased signal $\tilde{v} \in (W, S)$ in the first period. She delays her purchase at some price $p_1 \in (W, S_1)$ under information structure $\mathcal{I}$ (compared to $\mathcal{I}'$) if and only if knowing her true value *strictly* improves her expected utility in the second period; because $p_2 \sim D$ regardless of p_1, delay occurs if p_1 is smaller than but close to $v_1^{-1}(\tilde{v})$. We will demonstrate a positive measure of such $\tilde{v}$, so that the buyer delays purchase with strictly positive probability.

Now recall from Step 1 that a signal $\tilde{v} < S$ is only received when the true value also satisfies $v < S$. Because we assume $\Pi_{RSD} > \Pi^*$, Proposition 8 gives $W > \underline{v}$. Thus a positive measure of signals $\tilde{v} \in (W, S)$ is received when the true value v belongs to the interval $[\underline{v}, W)$.[29] We claim that for any such $\tilde{v}$, knowing the true value in the second period strictly benefits the buyer. This is because according to her expected value $\tilde{v} > W$, the buyer in the second period buys at some price p_2; but if she were informed that $v < W$, she would not buy at any price p_2 (which is at least W). This proves that by providing $\mathcal{I}$ rather than $\mathcal{I}'$, nature induces a positive probability of delay. By Lemma 6, we deduce that profit from the first buyer is less than Π_{RSD}.

Step 5: Proof for an arbitrary pricing strategy σ. Finally, we turn to prove the proposition 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 is close to H, which is its distribution if $\sigma = \sigma^D$.[30] The rest of the proof proceeds as in Step 4: a positive measure of signals $\tilde{v} \in (W, S)$ is received when the true value satisfies

[29] If $\Pi_{RSD} = \Pi^*$, then Proposition 8 implies $W = \underline{v} = p^*$ and Du's distribution is a mass-point at W. Information in the second period is irrelevant, because a buyer waiting till the second period always buys at price $p_2 = W = \underline{v}$.

[30] Consider nature choosing $\tilde{F}$ in the first period and doing nothing afterwards. The seller's profit from the first buyer can be written as $\mathbb{E}^w[w(1 - \tilde{F}(w)]$, where the random variable w equals $v_1(p_1)$ with probability $1 - \delta$ and it equals $\min\{v_1(p_1), p_2\}$ with probability δ (see (9)). The distribution of w must be close to D, otherwise nature could choose $\tilde{F}$ and damage profit from the first buyer. Since p_2 is approximately distributed according to D, we can derive as in the proof of Lemma 4 that $v_1(p_1)$ must be approximately distributed according to H.

$v < W$. For such $\tilde{v}$, full information in the second period is strictly valuable, and the buyer delays purchase if $v_1(p_1)$ is smaller than but close to $\tilde{v}$. By what we have shown, this occurs with strictly positive probability. But then Lemma 6 implies profit from the first buyer is bounded away from Π_{RSD} under $\mathcal{I}$, leading to a contradiction. The proof of Proposition 6 is complete. ■

Let us conclude by commenting on the assumption that Du's mechanism is uniquely optimal. Suppose this assumption fails, so that another point $\hat{S} > S$ satisfies (32). This means there are two different Du distributions D and $\hat{D}$, supported on $[W, S]$ and $[W, \hat{S}]$ respectively. On their supports, both of these distributions have density proportional to $\frac{1}{p}$ (see (31)). This observation allows us to write

$$\hat{D} = \alpha D + (1 - \alpha)E \tag{34}$$

with $\alpha \in (0, 1)$ is a scalar and E is a distribution supported on $[S, \hat{S}]$ (again with density proportional to $\frac{1}{p}$).

When such non-uniqueness occurs, the previous proof of Proposition 6 fails. Specifically, in Step 5, we are not able to deduce that σ is "close" to either σ^D or $\sigma^{\hat{D}}$. In fact, the following pricing strategy σ guarantees profit Π_{RSD} from the second buyer as well as from the first buyer, if nature chooses the information structure $\mathcal{I}$ in Step 2.

- The seller chooses a distribution of p_1 so that $v_1(p_1) \sim E$, which is supported on $[S, S']$. Here $v_1(p_1)$ is defined by the usual indifference condition $v_1 - p_1 = \delta \cdot \mathbb{E}^{p_2 \sim D}[\max\{v_1 - p_2, 0\}]$.
- Independent of the realized p_1, the seller draws $p_2 \sim D$, supported on $[W, S]$.

Because the price in the second period follows a Du distribution, the seller's profit from the second buyer is at least Π_{RSD}. For the first buyer, consider first the information structure $\mathcal{I}'$ as in Step 3, where nature reveals $\tilde{v} \sim F_W^B$ in the first period and no additional information afterwards. As shown in Footnote 30, the seller's profit from this buyer is $\mathbb{E}^w[w(1 - F_W^B(w)]$. This is as in the one-period model, where the seller charges price w and nature provides the Roesler-Szentes information structure.

Recall that w is a random variable that equals $v_1(p_1)$ with probability $1-\delta$ and $\min\{v_1(p_1), p_2\}$ with complementary probability. Because $v_1(p_1) \sim E$, whose support is strictly above the support of p_2, we deduce that $w \sim \delta D + (1 - \delta)E$. Thus, by (34), the distribution of w is a convex combination of D and $\hat{D}$ whenever $\delta \geq \alpha$. Since the seller ensures profit Π_{RSD} by using a random price distributed according to either D or $\hat{D}$, he does just as well by charging w. We have thus shown that profit from the first buyer is at least Π_{RSD} under information structure $\mathcal{I}'$.

Moreover, we claim that when nature provides $\mathcal{I}$ rather than $\mathcal{I}'$, no buyer delays her purchase. To see this, consider a buyer who purchases in the first period under $\mathcal{I}'$. By definition of v_1 and

the fact that $v_1 \sim E$, this means the buyer's signal $\tilde{v}$ in the first period satisfies $\tilde{v} \geq v_1(p_1) \geq S$. But then her true value v must also be at least S, as we showed in Step 1. Such a buyer purchases at any price $p_2 \in [W, S]$ regardless of any information in the second period. Thus, although nature promises future information under $\mathcal{I}$, this information does not improve the buyer's expected utility in the second period. Consequently the buyer's behavior under $\mathcal{I}$ is the same as under $\mathcal{I}'$, and profit under $\mathcal{I}$ is also equal to Π_{RSD}.

To summarize, we have constructed a pricing strategy σ such that if nature chooses the particular information structure $\mathcal{I}$ (for the first buyer), the seller's total profit is at least $(1+\delta)\Pi_{RSD}$. This explains why *our proof of* Proposition 6 requires the assumption that Du's mechanism is unique. We do not know whether the proposition is generally true without this assumption.[31]

[31] In other words, suppose S is not unique and suppose the seller uses the strategy σ constructed just now. We do not know whether nature can damage the seller's profit to be strictly lower than Π_{RSD} by choosing an information structure different from the $\mathcal{I}$ in our proof.