Informational Robustness in Intertemporal Pricing

Jonathan Libgober* and Xiaosheng Mu**

*University of Southern California
**Princeton University

August 11, 2020

ABSTRACT. We introduce a robust approach to study dynamic monopoly pricing of a durable good in the face of buyer learning. A buyer receives information about her willingness-to-pay for the seller's product over time, and decides when to make a one-time purchase. The seller does not know how the buyer learns, but commits to a pricing strategy to maximize profits against the worst-case information arrival process. We show that a constant price path delivers the robustly optimal profit, with profit and price both lower than under known values. Thus, under the robust objective, intertemporal incentives do not arise at the optimum, despite the possibility for information arrival to influence the timing of purchases. We delineate whether constant prices remain optimal (or not) when the seller seeks robustness against a subset of information arrival processes. As part of the analysis, we develop new techniques to study dynamic Bayesian persuasion.

KEYWORDS. Intertemporal Pricing, Dynamic Information Structures, Robustness, Information Design, Mechanism Design.

CONTACT. libgober@usc.edu, xmu@princeton.edu. We are particularly indebted to Drew Fudenberg for guidance and encouragement. We also thank Dirk Bergemann, Ben Brooks, Odilon Câmara, Gabriel Carroll, Yeon-Koo Che, Rahul Deb, Songzi Du, Kfir Eliaz, Ben Golub, Jerry Green, Johannes Hörner, Yuhta Ishii, Navin Kartik, Scott Kominers, Pablo Kurlat, Qingmin Liu, Eric Maskin, Luciano Pomatto, Tomasz Strzalecki, Guofu Tan, Juuso Toikka, and anonymous referees for comments and suggestions that improved the paper. Xiaosheng Mu acknowledges the hospitality of Columbia University and the Cowles Foundation at Yale University, which hosted him during parts of this research.

1. INTRODUCTION

This paper studies how the possibility of buyer learning influences dynamic monopoly pricing of a durable good. Following Stokey (1979), we consider a forward-looking buyer (she) who decides when to purchase a single unit of an object from a monopolist seller (he). Assuming that the buyer knows her value for the object and this value remains constant over time, the classic result of Stokey (1979) demonstrates that the seller's optimal pricing strategy is to use a constant price path. The key insight is that lowering future prices leads to increased sale to low-value buyers, but at the same time causes some high-value buyers to delay purchase. This trade-off results in a net loss in profit.

Contrary to this benchmark, we consider a buyer who does not initially know her value for the seller's product. Instead, she receives information about the product's worth or her idiosyncratic needs, and updates her beliefs about her value over time. Consider a consumer deciding whether and when to purchase a new car, as she worries that her current car may need some costly repairs. Not knowing exactly what kinds of repairs will be needed or how much inconvenience they will cause, she is imperfectly informed about her outside option, and therefore about her (net) lifetime discounted value for a new car. Furthermore, suppose that the buyer can learn about her current car's quality from her mechanic. By providing information, the mechanic influences the buyer's willingness-to-pay for the new car and thus the seller's profit as well.

With this example in mind, our theoretical question can be phrased as follows: Should the presence of the mechanic (i.e., information arrival) qualitatively change how the car seller sets prices? In particular, is it still optimal to keep prices constant, or does the possibility of buyer learning make intertemporal price discrimination profitable? In the latter case, are there general insights about the form of the optimal pricing policies?

1.1. An Example

The following simple example illustrates how optimal pricing is sensitive to information arrival. Suppose that a buyer has value v for the seller's product, with $\mathbb{P}[v = 4] = \frac{1}{4}$, $\mathbb{P}[v = 3] = \frac{1}{2}$, and $\mathbb{P}[v = 0] = \frac{1}{4}$. For simplicity, we assume that transaction can occur in one of two periods, with both parties discounting second period payoffs by a factor δ. First consider the "known-values" case studied in Stokey (1979), where the buyer knows v at the beginning of period 1. In this case, optimal dynamic pricing coincides with optimal static pricing; for the above distribution, a constant price of 3 yields the optimal expected profit of $\frac{9}{4}$, with no delayed purchase.

If the buyer learns about her value over time, then the seller would like to tailor the pricing strategy to the information arrival process. Suppose (the seller knows that) the buyer only knows whether or not $v = 4$ in the first period, but learns v perfectly in the second period. Consider the

pricing strategy which charges $4 - \delta$ in the first period and 3 in the second period. Given these prices, a buyer with value 4 purchases in period one (since she is indifferent), while a buyer with value 3 purchases in period two. This will lead to an expected profit of $1 + \frac{5}{4}\delta$. Thus, declining prices can facilitate price discrimination by inducing sale over time. Stokey's result fails here because, under this information arrival process, lowering the price to 3 in the second period leads to *additional* sales that would not have happened if the price were equal to 3 in the first period. We can further show this strategy is optimal for $\delta > \frac{4}{5}$; see Online Appendix F.1 for details.[1]

But if the seller anticipated a different information arrival process, then he might also price discriminate using an increasing price path. To illustrate, suppose that the buyer instead learns whether or not $v = 3$ in the first period and later learns v perfectly. Then charging 3 in the first period and 4 in the second period enables sale to occur in both periods, yielding an expected profit of $\frac{3}{2} + \delta$. This turns out to be the optimal strategy for $\delta > \frac{1}{2}$. Intuitively, information arrival leads to delayed purchase by $v = 4$ buyers who only learn their value in the second period, and increasing prices extract more surplus from these buyers than constant prices. This does not occur in Stokey's known-values setting.[2]

1.2. Model and Results

The preceding example illustrates the difficulty in providing a benchmark prediction regarding the seller's optimal pricing strategy in the presence of buyer learning.[3] In this paper, we focus on the case of a seller who *does not know* how the buyer receives information. We assume that the buyer's value is drawn from a commonly known distribution, and that the buyer learns her value over time according to some information arrival process (which she knows). However, unlike in the above example, the seller does not know the process and thus does not optimize against any particular one. Instead, he commits to a pricing strategy that *maximizes his profit guarantee* against all possible information arrival processes. In our baseline model, we consider the case of very rich informational uncertainty, where information in each period can depend on the seller's potentially randomized pricing strategy, as well as on realized prices up to and including that period. This gives the most cautious profit benchmark for the seller.

Returning to the car buyer, the seller may be unable to determine the tests the mechanic will

---

[1] When $\delta \leq \frac{4}{5}$, it is optimal to sell only in the first period at a price of 2, while shutting down sale in the second period by charging any price ≥ 3. This strategy yields a higher profit of 2 for low δ.

[2] Apart from buyer learning, other channels through which a departure from constant prices may be optimal include buyer budget constraint (Che and Gale (2000)), unequal discount factors between seller and buyer (Landsberger and Meilijson (1985)), and buyer value shocks (Deb (2014), Conlisk (1984), and Garrett (2016)).

[3] In more general dynamic mechanism design frameworks, Courty and Li (2000) and Pavan, Segal and Toikka (2014) have shown that the optimal selling mechanism depends sensitively on the process by which buyer value evolves, if the seller knows this process.

perform on the buyer's car, and hence be uncertain about the buyer's information (structure). The seller might even worry that the mechanic's goal is to prevent the buyer from purchasing a new car, which would need fewer repairs. More generally, if information relates to idiosyncratic tastes, the seller might have limited ability to anticipate how a particular buyer discovers her preference over time. He would therefore face uncertainty over how this buyer's expected value evolves, even if the ex-ante value distribution is commonly understood. This seller would prefer a pricing strategy that performs well in a variety of informational environments, rather than just one.

Our main result is that, under the robust objective, the seller optimally uses a constant price path, for any time horizon and discount factor. The optimal price and profit are both lower than the known-values case, due to the distinct seller objective in our model. However, sufficient uncertainty over the informational environment restores an important feature of Stokey (1979), namely that the dynamic problem can be reduced to optimal static pricing. Notice that this conclusion need not hold when the seller restricts to particular information arrival processes, as demonstrated by the above example.

Toward this result, we first argue that by always charging the optimal price in the single-period problem (which we describe below), the seller obtains the single-period profit guarantee even if information arrives *dynamically*. This claim is not immediate, since dynamic information can induce the buyer to delay purchase and potentially hurt discounted profit. What we show, however, is that the effect of delay on profit can be replicated by instead providing information in period 1 in a way that lowers the probability of sale. Intuitively, when the seller uses a constant price path, both seller profit and buyer payoff are determined by the discounted probability of sale given the buyer's true value. For any dynamic information structure, we can find a *static* information structure that makes only one purchase recommendation in period 1, with appropriate probabilities that maintain the total discounted probability of sale to each buyer type. In the formal argument, we additionally verify that the buyer is willing to follow such a recommendation. As a result, this static information structure is "outcome-equivalent" in terms of buyer and seller surplus. This equivalence, which we call the Replacement Lemma (Lemma 1), shows that constant prices are robust to dynamic information.[4] Our method of replacement turns out to be a useful tool more generally; we demonstrate this in various extensions, where modified versions of Lemma 1 are applied to reduce the dynamic problem to a static one.

The natural next question is: Why does the seller not benefit from intertemporal price discrimination, for example, by lowering prices over time? We recall that the classic intuition from the known-values case is based on the trade-off between selling to more buyers at a lower price (tomorrow) versus fewer buyers at a higher price (today). Given a fixed value distribution, this

---

[4]Although Lemma 1 is stated for the seller's worst-case profit, its proof shows that with constant prices, the replacement static information structure keeps both buyer and seller surplus the same.

trade-off is optimized by selling with probability 1 to all buyers with value above a certain level, where the "virtual value" equals zero. This optimum is implemented by a constant price path. As we have seen in Section 1.1, this standard intuition does not readily generalize in the presence of buyer learning, since information arrival changes the value distribution and influences the trade-off between earlier and later sales. However, we restore this intuition *under the robust objective*, by establishing a connection between our problem and the known-values case. We first observe that, with a single period, the worst-case information structure recommends the buyer to purchase if and only if her value is above a price-dependent threshold. The threshold is monotonic in the price: In fact, it has the property that the buyer's expected value, conditional on being below the threshold, exactly equals the price—such an information structure minimizes the probability of sale. This solution suggests that our single-period problem can be thought of as an *as-if known-values* problem, in which the prior value distribution is transformed to take into account the mapping from prices to thresholds (which reflect worst-case information).

Moving on to the dynamic problem, we generalize the threshold information structure to *threshold information arrival processes*, which inform the buyer in each period whether her value is above or below a threshold. Intuitively, threshold processes maximize the buyer's expected value when she is recommended to purchase (i.e., when her value is above the threshold), so as to minimize her purchase probabilities and thus the seller's profit. For any pricing strategy the seller uses, we exhibit a threshold process such that buyer behavior and seller profit coincide with the as-if known-values problem. This connection recovers the trade-off between selling to more buyers at a lower price versus fewer buyers at a higher price, when evaluating probabilities of sale according to the transformed value distribution (which does not change over time). As under known values, the seller in our problem does not gain from intertemporal price discrimination.

This analysis also reveals that a certain richness in the informational environment is necessary in order to reduce dynamic pricing to a static problem. While the solution to our baseline model provides a maximally cautious lower bound on the seller's profit, non-constant pricing may improve the profit guarantee if the seller is only concerned with a subset of possible information arrival processes. We illustrate this with two extensions of our main model. First, we consider cases where the buyer receives information infrequently, and show that a declining price path out-performs a constant price path. We then present a variant of our model with many buyers arriving over time, who share a common value and observe common signals. An increasing price path turns out to be optimal in the patient limit.

Our analysis echoes others in the robust mechanism design literature, which highlight that simple strategies can be optimal given sufficient uncertainty over the environment.[5] Constant

---

[5]See, e.g., Chung and Ely (2007), Frankel (2014), Carroll (2015, 2017) and Yamashita (2015). Among these, the closest

price paths are "simple" because the optimum can be achieved without knowing the buyer's arrival time or even the time horizon. Our result thus provides justification for firms to eschew sophisticated pricing strategies, even when consumer learning is significant. We find it reassuring that the worst-case information arrival process always takes the threshold form, as introduced above. Threshold information structures admit a natural interpretation as a "pass/fail" test, and have been studied in a variety of applications, from finance (Goldstein and Leitner (2018), Inostroza and Pavan (2019)) to political economy (Alonso and Câmara (2016)).[6] For our dynamic setting, we generalize static information structures that involve a single threshold (as in these papers) to dynamic information processes that involve multiple descending thresholds. So long as the seller seeks robustness against at least this class of processes, our analysis is unaffected.

Our study of informationally robust pricing is inspired in part by Bergemann, Brooks and Morris (2017), Du (2018) and Brooks and Du (2019). The goal of this line of research is to move away from specific assumptions about the informational environment, which may imply optimal mechanisms that depend sensitively on these assumptions. Relative to the existing work, we introduce *dynamic* informational robustness and demonstrate how constant pricing ensures robustness against *potential buyer delay*.[7] Prior literature has also studied pricing under uncertainty about the value distribution—see Bergemann and Schlag (2011), Handel and Misra (2014), Caldentey, Liu and Lobel (2016), Liu (2016), Chen and Farias (2018) and Carrasco et al. (2018).

Dynamic information arrival presents certain modeling challenges; in particular, there are potentially many ways to model the interaction between information and prices over time. Our main model allows information in each period to adapt to prices up to and including that period. We view this generality as desirable, since it delivers the most cautious profit benchmark. For example, the car seller mentioned above might worry that the buyer's mechanic observes the new car's price before deciding what to reveal about the old car's breakdown risk. In practice, a variety of channels may lead to price-dependent information; see further discussion in Section 2.3. A less cautious benchmark would have been to disallow such price-dependence. This alternative model was studied by Du (2018) for a single period, building on the earlier work of Roesler and Szentes (2017). In Section 5, we study a dynamic version of that model and show that a *randomization*

---

to our paper is Carroll (2017), who shows how uncertainty over the correlation between a buyer's demand for different goods leads the seller to price the goods independently. The general link between dynamic allocations and multi-dimensional screening has been noted in the literature. While it is interesting that we obtain a result similar to Carroll (2017), our focus on information arrival and single-object purchase distinguishes from that work.

[6]Also related are Bergemann and Wambach (2015) and Li and Shi (2017). These papers discuss that threshold information might arise via comparisons to past products for which buyer values are known. Threshold processes might also arise if the product has several attributes that are sequentially revealed to the buyer, who has lexicographic preference over these attributes.

[7]Most papers in robust mechanism design focus on static settings. As far as we are aware, Chassang (2013) and Penta (2015) are among the few papers that study a dynamic robust objective.

over constant price paths delivers the robustly optimal profit.

Methodologically, our analysis contains certain technical innovations that may be applicable to other problems, particularly those that involve Bayesian persuasion. The connection to the persuasion literature (Kamenica and Gentzkow (2011) and many that follow) arises since our seller is worried about an "adversarial nature" who attempts to persuade buyers not to purchase the product. Viewed from this perspective, our results provide a characterization of optimal persuasion (i.e., worst-case information structure) by nature against a given pricing strategy. In particular, our Proposition 3 shows it is without loss to restrict attention to threshold information arrival processes. This is a dynamic version of the optimality of interval persuasion previously established for static models, such as in Kolotilin (2015) and Dworczak and Martini (2019). We also suspect that our Replacement Lemma may have broader relevance to dynamic Bayesian persuasion, as it suggests that certain instances of such problems may admit static solutions.

2. MODEL

Our baseline model adds buyer learning to an otherwise standard dynamic pricing setting. A seller (he) sells a durable good at times $t = 1, 2, \ldots, T$, where $T \leq \infty$. For now, we assume there is a single buyer (she), present at time $t = 1$, who can delay purchase to any later time; the case where multiple buyers arrive over time will be discussed later. Both seller and buyer have discount factor δ. The product is costless for the seller to produce,[8] while the buyer has unit demand. The buyer has (undiscounted) lifetime value v from purchasing the object, where v is drawn from a distribution F and fixed over time. The prior distribution F is common knowledge, with support $V \subset \mathbb{R}_+$ and $0 < \mathbb{E}[v] < \infty$. For expositional ease, we assume that F is a continuous distribution with minimum value $\underline{v}$; we explain how our results extend to discrete distributions in Appendix A. As in the mechanic story above, the value v can also be interpreted as the buyer's net value for the seller's product relative to an outside option that she learns about.

At time 0, the seller commits to a pricing strategy σ, which is a distribution over possible price paths $p^T = (p_t)_{t=1}^T \in \mathbb{R}_+^T$.[9] The buyer decides when to purchase based on her knowledge of the seller's strategy, the price in that period, as well as her belief about her value and what she expects to learn about her value in the future.[10] The next subsection formalizes the learning process. A

---

[8]Introducing a cost of c per unit does not change the results for our main model. It is as if the prior distribution F were "shifted down" by c, and the buyer might have a negative value. The pressed distribution G in Definition 1 below would simply be shifted down by c as well.

[9]The commitment assumption is frequently made in the intertemporal pricing literature. In our setting, dropping commitment would introduce further difficulties related to formalizing learning under ambiguity; see Epstein and Schneider (2007) and Riedel (2009).

[10]We assume that the buyer knows her information arrival process, and is Bayesian about what information will be received in the future. However, our analysis is unchanged if the buyer also faces uncertainty (and is maxmin) over

buyer who never purchases the object obtains a payoff of 0.

2.1. Dynamic Information Structures

The buyer does not directly know v; instead, she learns about it through signals that arrive over time, via some information structure. To be precise, a *dynamic information structure/information (arrival) process* $\mathcal{I}$ consists of:

- A set of possible signals for every time $t \geq 1$, i.e., a sequence of sets $(S_t)_{t=1}^T$, and

- Probability distributions given by $I_t : V \times S^{t-1} \times P^t \to \Delta(S_t)$, for all t with $1 \leq t \leq T$.

Above, $S^{t-1} = \prod_{\tau=1}^{t-1} S_\tau$ denotes the set of possible past signal realizations, and $P^t := \mathbb{R}_+^t$ represents the set of possible past and current prices. To be fully rigorous, there should be a σ-field associated with each S_t, and the mappings I_t are required to be measurable. We will however omit these technical details, which do not affect the analysis.

To interpret the above definition, note that the distribution of the signal s_t at time t could depend on the buyer's true value $v \in V$, the history of her previous signal realizations $s^{t-1} = (s_1, \ldots, s_{t-1}) \in S^{t-1}$, as well as the history of *all previous and current prices* $p^t = (p_1, p_2, \ldots, p_t) \in P^t$. The possibility for information to flexibly depend on realized prices distinguishes our model from other papers using the robust approach, and we discuss this important assumption more thoroughly in Section 2.3 below. For now, we simply point out that if the seller were to use a deterministic price path, our definition would reduce to the usual definition that signal s_t occurs with probability $I_t(s_t \mid v, s^{t-1})$. In that case we could omit the dependence on realized prices since there is only one possible realization. As we discuss later, allowing for price-dependent information only has bite when the seller randomizes.[11]

2.2. Seller's Objective

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

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

---

future information, so long as she can interpret signals in the current period. This extension is discussed in Online Appendix E. In this sense, we do not impose extra rationality of the buyer beyond what is typically assumed in static robust mechanism design.

[11]Since a deterministic (constant) price path is optimal in our main model, an alternative model where information can further condition on future price realizations would yield the same result.

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 allow the stopping time τ 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 process 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. It will follow from our analysis that when the prior distribution F is continuous, $\sup \inf$ is achieved as $\max \min$.[12] Breaking indifference in favor of the seller would not change our results, but would add cumbersome details due to $\max \min$ not being achieved.

2.3. Discussion

Relation to known-values. Our main model assumes that both parties start off with the same prior about the buyer's value. In Section 4.2, we show that our constant price path result is maintained when it is common knowledge that the buyer has a more informed prior. An extreme case of this extension is when the buyer perfectly knows her value (and the seller knows that), which corresponds to a discrete-time version of Stokey (1979). In this sense, our result extended to the setting in Section 4.2 is a strict generalization of Stokey (1979).

Informational versus distributional uncertainty. We focus on the study of seller uncertainty regarding buyer learning, and for this reason shut down any uncertainty about the prior distribution F. This captures settings where heterogeneity in buyers' willingness-to-pay is primarily due to idiosyncratic tastes that are discovered over time. However, our framework can be extended to accommodate aggregate value uncertainty. In Online Appendix D, we discuss how the presence of distributional uncertainty–on top of informational uncertainty–would influence our analysis. We show that for any set of possible value distributions, those distributions that are worst for profit are minimal with respect to second-order stochastic dominance (Theorem 3). In particular, if the seller does not know the value distribution but knows its mean and range, then the worst-case distribution is supported on the extreme values, and the seller charges the optimal constant price (given in our Proposition 1) against this fixed distribution.[13] More generally, Theorem 4 shows

---

[12] This result may fail if there is a large mass point at $\underline{v}$. See the proof of Proposition 1 in Appendix A for details.

[13] This relates to Carrasco et al. (2018), which considers a seller who does not know the distribution of the buyer's value, but knows some of its moments.

that the worst-case distribution exists under regularity conditions.

Price-dependent information. A key assumption in our model is that the seller is maximally cautious, in the sense that he does not rule out any information process for the buyer. In the spirit of delivering the most cautious benchmark, our baseline model further allows for information to be *price-dependent*.[14] In reality, such price-dependence could occur through a number of channels: For instance, if advertisements are displayed more prominently depending on price, if reviewers consider price when deciding which products to write about, or if buyers are rationally inattentive and choose information based on the price.

Including price-dependent information additionally provides technical convenience in that the ability of nature to respond to realized prices eliminates the seller's incentive to randomize, despite the maxmin objective.[15] With the restriction to deterministic prices, we can describe the solution to our problem (and its intuition) in a way similar to the known-values case of Stokey (1979). As we show in Section 6, developing this analogy also helps us understand the boundaries of the constant price path result.

Interestingly, when the seller's uncertainty is restricted to *price-independent* information, a randomization over constant price paths becomes optimal even though the analogy to the known-values case is lost. We discuss this extension in Section 5.

3. ONE-PERIOD ANALYSIS

We start by analyzing the one-period problem. To solve this problem, we will define a transformed distribution of the prior F. This transformation uses our assumption that F is continuous, with minimum value $\underline{v}$. Our main 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, its "pressed version" G is another distribution defined as follows. For $y > \underline{v}$, let $L(y) = \mathbb{E}[v \mid v \leq y]$ denote the expected value (under F) conditional on the value not exceeding y.[16] Then $G(\cdot) = F(L^{-1}(\cdot))$ is the distribution of $L(y)$ when y is drawn according to 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

---

[14] As discussed in a decision theory framework by Ke and Zhang (2019), any assumption on how information interacts with prices is related to the seller's subjective model of the *timing of nature's moves relative to his own randomization*. In our dynamic setting, there are multiple ways one could model the timing of nature's moves. Our main model takes the most pessimistic perspective that nature moves in each period, after the seller's randomization.

[15] This is straightforward to see in one period, but also true in many periods as we show.

[16] The conditional expected value $L(y)$ is closely related to the notion of *conditional value at risk/expected shortfall* in mathematical finance, see e.g. Ma and Wong (2010).

structure is outcome-equivalent to another that directly recommends one of two actions: to purchase the good or not. Given this simplification, the worst-case information structure must have the following property: As long as the buyer is recommended to purchase with positive probability, the buyer who is recommended *not* to purchase has expected value exactly p. Otherwise nature could adjust its recommendation to further decrease the probability of sale.

Moreover, given that a buyer who does not buy has fixed expected value (in our case, p), one can show that a *threshold information structure* maximizes the probability of this recommendation (see e.g. Kolotilin (2015)). In a threshold information structure, the buyer is told whether her value is above or below a certain threshold. By the above definition of G, this threshold must be $L^{-1}(p) = F^{-1}(G(p))$, making $1 - G(p)$ the probability of sale.

These observations 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:*

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

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

The optimization problem (1) is exactly analogous to the seller's problem under known values. If the buyer knew her value, the seller would maximize $p(1 - F(p))$. In our setting with informational uncertainty, the difference is that the pressed distribution G takes the place of F. Our analysis in the next section reveals how this analogy can be extended to the dynamic model.

The following example illustrates our transformation:

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 $\frac{1}{4}$. In response, nature chooses an information structure that tells the buyer whether or not $v > \frac{1}{2}$.*

We mention that 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 $\frac{1}{2}$, since such a buyer will purchase in any event.

In this example, relative to the known-values case, the seller charges a lower price and obtains a lower profit under informational uncertainty. The following proposition shows this comparison is general:

Proposition 2. *For any distribution F, let $\hat{p}$ be an optimal monopoly price under known values:*

$$\hat{p} \in \underset{p}{\operatorname{argmax}} \, p(1 - F(p)), \tag{2}$$

and let $\hat{\Pi} = \hat{p}(1 - F(\hat{p}))$ be the corresponding profit. Then any maxmin optimal price p^* satisfies $p^* \leq \hat{p}$, and the maxmin profit satisfies $\Pi^* \leq \hat{\Pi}$. Equality holds only if $p^* = \hat{p} = \underline{v}$.

Appealing to the "F-to-G transformation" allows us to derive further results on how the seller's maxmin profit varies with the prior distribution. Intuitively, greater variation in the prior value distribution gives rise to greater uncertainty about what the buyer may learn. We would thus expect that under the robust objective, the seller is worse off if F decreases with respect to second-order stochastic dominance. In Online Appendix D, we show this is indeed the case by demonstrating that second-order stochastic dominance in F is equivalent to first-order stochastic dominance in the pressed distribution G (Lemma 7).

4. MAIN RESULTS

With multiple periods, the following is our main result:

Theorem 1. *The seller's maxmin optimal profit is Π^*, given any selling horizon T and discount factor δ. This maxmin profit is achievable by a constant price path of p^* charged in every period.*

As we saw through the example in Section 1.1, if the seller knew the information process, he would want to adapt his pricing strategy to this particular process, in order to facilitate price discrimination. Nonetheless, optimal prices for a fixed information process could increase or decrease over time, depending on how one specifies the process.

In contrast, Theorem 1 suggests that when facing uncertainty over buyer learning and adopting a robust objective, the seller is best off committing to the simple strategy of a constant price. Thus, by using the robust approach, we are able to restore the benchmark prediction of optimal constant pricing (Stokey (1979)) in a setting with buyer learning.

The underlying mechanism for our result is more involved than the case of known values. Indeed, information arrival may cause a buyer to delay purchase when facing a constant price path—but we show this does not occur *in the worst case*. One may worry that constant price paths perform well because they guard against some contrived information processes. As we explain later in this section, this is not a concern for our problem. Our result is unchanged so long as the seller seeks robustness against the intuitive class of "threshold information processes." Finally, while we believe it is of theoretical interest to generalize the classic result of Stokey (1979),

perhaps more important are the assumptions that give rise to it.[17] In this sense, our constant price path result provides a benchmark to understand if and when restrictions on the informational environment can lead to price dynamics. Later in Section 6, we present some results of this form, where dynamic pricing out-performs constant pricing.

4.1. Proof Sketch of Theorem 1

Here we outline the arguments we use to prove Theorem 1; the detailed proofs can be found in Appendix A. Our proof separately establishes a lower-bound and an upper-bound on the seller's profit guarantee. For the lower-bound, we argue that by using a constant price path, the seller obtains at least Π^* from the buyer under any information process. This follows from our Replacement Lemma, which shows that for non-decreasing prices, any dynamic information structure can be replaced with a static one while weakly lowering profit. We then demonstrate a matching upper-bound: No matter how the seller sets prices, nature can hold profit to at most Π^*. This part of the argument takes advantage of the intuition from the one-period analysis and generalizes the threshold information structure appropriately to the dynamic setting.

Note that the upper-bound is by itself sufficient to imply that the seller's maxmin profit is Π^*, since he can choose to sell only in the first period and achieve the lower-bound. In fact, by Lemma 1 below, any increasing price path with $p_1 = p^*$ would guarantee this profit. But the constant price path of p^* has an additional advantage of being stationary and thus robust to the possibility of multiple buyers arriving over time. In Appendix A.3.4, we establish the *unique* optimality of constant pricing in such an environment.

Below we provide some details of the two parts of our proof, respectively.

4.1.1. Lower-bound

Under known values, a 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. A priori, such delay may hurt the seller's profit. Nonetheless, in the following lemma, we show that against a non-decreasing price path (among others), nature cannot hurt the seller more than providing information *only* in the first period.

Lemma 1 (Replacement Lemma). *Suppose that the seller uses a deterministic price path $(p_t)_{t=1}^T$ satisfying $p_1 \leq p_t, \forall t$. Then the seller's profit is minimized by an information structure that only provides information in the first period.*

---

[17] Although our main model does not nest the known-values setting, we present in Section 4.2 an extension that embeds our main model as well as known values. A constant price path remains optimal in that extension, thereby generalizing both Stokey's result and Theorem 1.

We call this result the "Replacement Lemma" because it shows that when prices increase over time, any dynamic information structure can be replaced with a static information structure that weakly decreases the seller's profit. Since delay does not occur under a static information structure and a non-decreasing price path, our previous one-period analysis shows that the seller obtains profit at least $p_1(1 - G(p_1))$, which equals Π^* when choosing $p_1 = p^*$. We note that the no-delay result is straightforward in settings without buyer learning (Stokey (1979), Riley and Zeckhauser (1983)), where delayed purchase at the same price can only hurt the buyer due to discounting. In our setting, a specific information arrival process can encourage delay even under constant prices. But with a rich set of possible information arrival processes, our analysis shows that assuming no-delay is without loss for understanding the range of payoff outcomes that can arise.

To construct such a replacement, we view the original dynamic information structure as providing recommendations to the buyer to purchase or not at different times. Whenever she was recommended to purchase in period t, in the replacement information structure we have nature recommend that she purchase in period 1 with probability δ^{t-1}. In other words, we "push and discount" nature's recommendation to period 1. The key technical step is to show that the buyer is still willing to follow nature's recommendation; we do this by using her incentive compatibility under the *original* information structure.[18] Once this is proved, it follows that the discounted probability of sale is unchanged, so that profit decreases (since prices are higher in future periods).

Looking ahead, we mention that similar methods of replacement play an important role for analyzing two variations of our model. See Lemma 3 and Lemma 4 in later sections.

4.1.2. Upper-bound

The second half of the proof of Theorem 1 involves constructing information processes that hold the seller's profit to Π^*, for any given pricing strategy. We look for information processes within the following class:

Definition 2. *Suppose the prior value distribution F is continuous.[19] A (descending) threshold*

---

[18]We provide an intuitive explanation. On the one hand, a buyer who is recommended to purchase in the replacement information structure has expected value at least p_1, since she was originally recommended to purchase at some price $p_t \geq p_1$. On the other hand, we need to show that a buyer who is recommended not to purchase has expected value at most p_1. Suppose expected buyer surplus under the original information structure was U. Since purchasing in period 1 regardless of the signal yields surplus $\mathbb{E}[v] - p_1$, we have $U \geq \mathbb{E}[v] - p_1$. Under the replacement, buyers recommended to purchase face the same discounted probability of purchase (δ^{t-1}) as in the original information structure, and the expected value conditional on purchasing is unchanged as well. Since they now pay $p_1 \leq p_t$, these buyers generate surplus $U' \geq U$, implying $U' \geq \mathbb{E}[v] - p_1$. Now observe that $\mathbb{E}[v] - p_1$ is the total surplus if *every* buyer purchases at price p_1. Thus, the remaining buyers who are not recommended to purchase would generate (weakly) negative surplus $\mathbb{E}[v] - p_1 - U'$ if they were to purchase at price p_1. Hence it is optimal for them to follow the recommendation and not purchase.

[19]In case F has atoms, we provide a generalized definition at the beginning of Appendix A.

information process involves a descending sequence of (possibly randomized) thresholds $x_1 \geq x_2 \geq \cdots \geq x_T$, where each x_t is measurable with respect to realized prices $p_1, \ldots, p_t$.[20] *Under this process, in each period t the buyer is told whether or not $v > x_t$.*

This generalizes the threshold information structures we introduced in Section 3, when studying the single-period problem.

By appealing to economic intuitions for our environment, we will construct a particular threshold information process that allows us to prove the following lemma:

Lemma 2 (Profit Upper-bound). *For any pricing strategy, there exists a threshold information process and a corresponding optimal stopping time that lead to profit $\leq \Pi^*$.*

To explain our construction, we assume for simplicity that the seller charges a *deterministic* price path $(p_t)_{t=1}^T$. If the buyer knew her value, then we could 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 optimally buys in period t_j whenever her value is $v \in [w_{t_j}, w_{t_{j-1}}]$. Here w_{t_j} is defined by the indifference condition $w_{t_j} - p_{t_j} = \delta^{t_{j+1}-t_j} \cdot (w_{t_j} - p_{t_{j+1}})$, and the fact that higher-value buyers purchase earlier is the well-known "sorting property" established for example in Stokey (1979). This implies that under known values, the object would be sold with probability $F(w_{t_{j-1}}) - F(w_{t_j})$ in period t_j.

In our setting, we find a threshold information process such that 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 thresholds defining the process are given as follows: In each period t_j, the buyer is told whether or not her value is in the lowest $G(w_{t_j})$-percentile, so that the threshold is $x_{t_j} = L^{-1}(w_{t_j})$ (see Definition 1). In other periods—that is, between any period t_j and t_{j+1}—no information is revealed, and $x_t = x_{t_j}$ at these periods. As in the one-period analysis, these 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, as future information can only improve her future payoffs. On the other hand, a buyer whose value is above the threshold does not expect to receive further information, and hence purchases immediately.[21]

The above observations show that $G(w_{t_{j-1}}) - G(w_{t_j})$ is the probability of sale in period t_j.

---

[20]We thank an anonymous referee for suggesting the terminology of "threshold information."

[21]We mention that the analysis is unchanged if any buyer with value above the current threshold perfectly learns her true value, since she purchases regardless. In this sense, the threshold information process we construct is outcome-equivalent to one where higher-value buyers discover their true values earlier.

We can then compute the seller's profit as follows:

$$
\begin{aligned}
\Pi &= \sum_{j \geq 1} \delta^{t_j - 1} p_{t_j} \cdot \big(G(w_{t_{j-1}}) - G(w_{t_j}) \big) \\
&= \sum_{j \geq 1} (\delta^{t_j - 1} p_{t_j} - \delta^{t_{j+1} - 1} p_{t_{j+1}}) \cdot (1 - G(w_{t_j})) \\
&= \sum_{j \geq 1} (\delta^{t_j - 1} - \delta^{t_{j+1} - 1}) w_{t_j} \cdot (1 - G(w_{t_j})) \\
&\leq \delta^{t_1 - 1} \cdot \Pi^*,
\end{aligned}
\tag{3}
$$

where we assumed $T = \infty$ for ease of illustration. The second line above is by Abel summation,[22] the third line uses type w_{t_j}'s indifference between buying in period t_j or t_{j+1}, and the last inequality holds because $w_{t_j}(1 - G(w_{t_j})) \leq \Pi^*$ for each j. This proves Lemma 2 when prices are deterministic.

To summarize, the key idea is that threshold information processes can force the same trade-off between later and earlier sale, just as under known values. Using the pressed distribution, nature can set the thresholds appropriately so that lowering prices in the future leads to (sufficiently) less sale in the current period. Thus the seller does not benefit from intertemporal price discrimination, and the single-period optimal profit guarantee remains an upper-bound in the dynamic setting.

This same intuition applies when prices are random, although in this case the indifference types w_{t_j} will be random variables and additional care is required. Formally, we define v_t to be the value type that is indifferent (under known values) between purchasing in period t and continuing to future periods. We then let $w_t = \min\{v_1, \ldots, v_t\}$ to denote the "binding indifference type", so that a buyer with known value in $(w_t, w_{t-1}]$ would optimally purchase in period t. Thus, the probability of sale in period t under known values is given by the random variable $F(w_{t-1}) - F(w_t)$. Similar to the above, we construct a threshold information process with thresholds $L^{-1}(w_t) = F^{-1}(G(w_t))$, and show that it yields probability of sale $G(w_{t-1}) - G(w_t)$. This then enables us to write the seller's discounted profit as a convex sum of one-period profits, generalizing the profit upper-bound in (3). Details of this proof are left to Appendix A.

4.1.3. Worst-case is Threshold Process

Having established the lower-bound as well as the upper-bound, we have completed the proof of Theorem 1. However, we note that the threshold information process constructed in the above

---

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

upper-bound argument is just one particular process that holds profit below Π^*. Example 2 in Appendix A.3.3 shows that this process does not in general deliver the worst case profit.

Despite this, we show below that the worst case dynamic information structure always falls within the class of threshold processes. This generalizes the optimality of interval persuasion (Kolotilin (2015)) to a dynamic setting.

Proposition 3. *Given any pricing strategy σ, there exists a (descending) threshold information process that minimizes the seller's profit.*

The basic intuition is familiar from the one-period analysis: To hurt the seller, it is best to *maximize* the buyer's expected value when she is recommended to purchase, so as to *minimize* the probability of such an event. This is achieved by providing threshold information. That said, accommodating dynamics introduces a new challenge since nature needs to trade off minimizing the probabilities of sale in different periods. Our proof in the appendix gets around this issue by replacing an arbitrary information process with a threshold one, such that the buyer's purchase times are stochastically later. Note that if the buyer delays purchase, incentive compatibility requires her expected payoff to increase, but social surplus must decrease due to discounting. We conclude that the seller's profit must be lower under the threshold information process.

Proposition 3 tells us that a seller concerned about the worst case need only worry about the simple class of threshold processes. Nonetheless, it remains challenging to solve for the exact worst-case (threshold) information process against any given pricing strategy.[23] This is due to difficulties with determining buyer optimal stopping under arbitrary prices and information. For this reason, we proved the upper-bound in Theorem 1 without computing the worst-case profit, but rather constructed a particular process that allows for easy computation of profit. Our approach takes advantage of the analogy to the known-values problem and delineates the intuitions for why dynamic pricing is not profitable for the seller.

4.2. Buyer with More Informed Prior

We now illustrate how a small extension of our model fully nests both our baseline model with complete informational uncertainty and the known-values case studied in Stokey (1979). In this extension we strictly generalize the constant price path result from the known-values case.

Specifically, we augment the model from Section 2 as follows: Suppose that, at time 0 and before the seller chooses a pricing strategy, the buyer observes a signal $\underline{s}$ drawn according to some *initial information structure $\mathcal{H}$.* We suppose this information structure (i.e., how $\underline{s}$ is distributed

---

[23]With non-decreasing prices, the worst case has been characterized in the Replacement Lemma. Solving for the exact worst-case process against a decreasing price path is an open question for future work.

given v) is common knowledge, but the seller does not observe the realization of $\underline{s}$. Thus, in this model, the buyer begins with weakly better information about her value than the seller. All other aspects of the main model are maintained, except that we allow nature to provide information conditional on $\underline{s}$.[24]

To study this extension, we let $F_{\underline{s}}$ be the buyer's posterior value distribution upon observing signal $\underline{s}$. The same analysis shows that for this "prior" value distribution, the worst-case static information structure involves a threshold. Hence, if we let $G_{\underline{s}}$ be the pressed distribution of $F_{\underline{s}}$, we have 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:*

$$p_{\mathcal{H}}^* \in \operatorname*{argmax}_{p} p(1 - \mathbb{E}[G_{\underline{s}}(p)]), \tag{4}$$

where the expectation is taken with respect to different realizations of the initial signal $\underline{s}$.

Denote the resulting one-period optimal profit by $\Pi_{\mathcal{H}}^*$. We can generalize Theorem 1 to this setting by following the same arguments as outlined in Lemma 1 and Lemma 2:

Theorem 1'. *Suppose that the buyer observes initial information structure $\mathcal{H}$. Then the seller's maxmin optimal profit is $\Pi_{\mathcal{H}}^*$, given any selling horizon T and discount factor δ. This maxmin profit is achievable by a constant price of $p_{\mathcal{H}}^*$.*

We return to our main model if $\mathcal{H}$ is uninformative, in which case the expectation in (4) is simply $G(p)$. On the other hand, if the initial signal $\underline{s}$ reveals the value v perfectly, then $G_{\underline{s}}(p)$ is equal to 1 if $p \geq v$ and 0 if $p < v$. In that case, $\mathbb{E}[G_{\underline{s}}(p)] = F(p)$, and we return to the known-values case of Stokey (1979) (although in this case Lemma 1 and Lemma 2 are vacuously true).

5. PRICE-INDEPENDENT INFORMATION

Our baseline model allows nature to provide information depending on all realized prices, delivering the most pessimistic profit guarantee for the seller. In this section, we study optimal pricing when information does not vary with realized prices.[25] With a single period, this modification

---

[24]It is equivalent to think of the initial information structure $\mathcal{H}$ as a constraint on nature's information choice in our baseline model. That is, we can view this extension as the seller seeking robustness only against those information processes such that the signal in period 1 is more informative than $\mathcal{H}$ in the sense of Blackwell (1953).

[25]Ruling out price-dependence requires less caution from the seller than our baseline model, but there may be cases where this is justified. For instance, suppose the seller is confident that consumers will learn about the product from product reviewers who follow the seller closely. In that case the seller may think that whether he charges 99 dollars or 89 dollars will not impact the amount of information buyers have access to.

connects to the recent models of Roesler and Szentes (2017) and Du (2018). The setups in these papers differ from our formulation, but their results imply the solution to this variant of our one-period model (with price-independent information). We describe these papers and results in more detail below.

We then analyze the dynamic version of this model and find that the seller can achieve the optimal profit guarantee by *randomizing over constant price paths*. In particular, we establish another version of the Replacement Lemma (Lemma 3 below), which shows that against randomized constant prices, nature cannot hurt the seller more than providing information only in period 1. Thus, by charging constant prices, the seller reduces the space of uncertainty from dynamic information processes to static information structures. Since the buyer with a fixed value distribution (induced by the static information structure) does not delay purchase when facing a constant price path, the seller again reduces the dynamic selling problem to a static problem.

5.1. Static Model: Relation to Roesler-Szentes (2017) and Du (2018)

This subsection describes a variant of our one-period model in which information does not depend on the realized price. We will also explain how this alternative model connects to the recent papers of Roesler and Szentes (2017) and Du (2018).

Formally, consider the following *zero-sum game* between the seller and nature, where nature seeks to minimize the seller's profit:

- The seller's strategy is a distribution of prices $\sigma \in \Delta(P)$, where $P = \mathbb{R}_+$.

- Nature's strategy is an information structure, consisting of a signal set S and a function $I : V \to \Delta(S)$, where $I(v)$ is the distribution over signals observed by the buyer with true value v.

- Finally, the buyer observes s and p, and purchases if and only if $\mathbb{E}[v \mid s] > p$.

The crucial change from our main model is that nature's choice of information structure is now *independent of the realized price* p. To reflect this, we describe the distribution over signals by the function $I : V \to \Delta(S)$, rather than $I : V \times P \to \Delta(S)$ as we did before.

This price-independent version of our one-period model relates to Roesler and Szentes (2017), who study the following interaction between a buyer and a seller. As in our model, the buyer's value is drawn from a commonly known prior distribution, and the buyer (as well as the seller) does not initially know her value. But in Roesler and Szentes (2017), the buyer chooses an information

19

structure, while the seller sets a profit-maximizing price in response to this choice.[26] Finally, the buyer observes her signal and decides whether or not to purchase. Note that in this model, the buyer seeks to maximize her own expected payoff and there is no adversarial nature. Nonetheless, this is connected to our maxmin framework because Roesler and Szentes (2017) show that the "buyer-optimal information structure" also minimizes the seller's profit (their Corollary 1).[27]

Phrased in the context of the above zero-sum game, the results of Roesler and Szentes (2017) characterize a *minimax* information structure that holds the seller's profit to the lowest level. Specifically, the *Roesler-Szentes information structure* induces posterior expected values with the following distribution:

$$F_W^B(s) = \begin{cases} 0 & s < W \\ 1 - \frac{W}{s} & s \in [W, B) \\ 1 & s \geq B \end{cases} \tag{5}$$

where W and B are numbers that depend on the prior distribution F. As Roesler and Szentes (2017) show, these numbers are such that the posterior distribution F_W^B is a mean-preserving contraction of the prior F, and that W is smallest possible subject to this constraint (see Online Appendix B.1 for further details). When nature (or the buyer) chooses this information structure, the seller's profit is bounded above by W regardless of his pricing strategy.[28]

On the other hand, Du (2018) shows that the seller can also guarantee profit W in the above zero-sum game. This optimal profit guarantee is achieved if the seller charges a *random price* with the following c.d.f.:[29]

$$D(p) = \begin{cases} 0 & p < W \\ \frac{\log \frac{p}{W}}{\log \frac{S}{W}} & p \in [W, S) \\ 1 & p \geq S \end{cases} \tag{6}$$

The number W is the same as in the Roesler-Szentes information structure; the number S belongs to the interval $[W, B]$, and is derived in Online Appendix B.1.[30] In Online Appendix F.3, we further

---

[26]There is no need to randomize in Roesler and Szentes (2017), since the seller moves after the information structure is chosen. In related work, Terstiege and Wasser (2019) consider optimal buyer information acquisition that is robust to potentially more information provided by the seller. Condorelli and Szentes (2018) study the problem where the buyer chooses her optimal value distribution.

[27]In Online Appendix F.2, we provide a related Bayesian interpretation of our model and results.

[28]As Roesler and Szentes (2017) point out, this posterior value distribution is the least amount of information (in terms of SOSD) that holds profit below W, but it is in general not the unique one. This will not affect our analysis.

[29]This construction generalizes Proposition 5 in Carrasco et al. (2018), who focus on prior distributions F with binary support.

[30]One difference from Du (2018) is that he allows the seller to use general mechanisms that prescribe allocation probabilities based on buyer reports. However, Du (2018) observes that with a single agent, the same outcome (i.e., profile of interim purchase probabilities) can be implemented using a randomization over posted prices, which

show that the seller's optimal strategy is unique for generic prior distributions F.

Taken together, the results from these two papers tell us that W is the seller's maxmin profit in the one-period problem with price-independent information. For future reference, we denote W by Π_{RSD}, after the authors of those papers. It is clear that Π_{RSD} is weakly larger than Π^*, and in Online Appendix F.4 we characterize when the comparison is strict.

5.2. Dynamic Model with Price-Independent Information

We now present a dynamic version of the above model with price-independent information. We re-define a dynamic information arrival process $\mathcal{I}$ to be a sequence of signal sets $(S_t)_{t=1}^T$, and probability distributions given by $I_t : V \times S^{t-1} \to \Delta(S_t)$, for all t with $1 \leq t \leq T$.

Similar to the preceding subsection, the interaction we study is a zero-sum game between the seller and nature:

- The seller chooses a pricing strategy $\sigma \in \Delta(P^T)$.

- Nature chooses an information process $\mathcal{I}$.

- Given σ and $\mathcal{I}$, the buyer chooses an optimal stopping time.

For this model, we characterize the seller's optimal pricing strategy and nature's worst-case information structure in the following theorem:

Theorem 2. *Suppose that information is independent of realized prices. The seller's maxmin optimal profit is Π_{RSD}, given any selling horizon T and discount factor δ. The seller can achieve this by randomizing over constant price paths drawn from Du's price distribution $D(p)$ in (6). Nature can force this profit upper-bound by only providing the Roesler-Szentes information structure to the buyer in period 1.*

It is not difficult to understand nature's information choice. By providing the static Roesler-Szentes information structure, nature makes the buyer "know her value" to be drawn from the distribution F_W^B. By the result of Stokey (1979), this holds profit below Π_{RSD}. Thus, the upper-bound on seller profit is immediate in this model, unlike in our baseline model.

The more striking feature of Theorem 2 is that the seller can guarantee Π_{RSD} by randomizing over constant price paths. This is proved via a generalization of the earlier Replacement Lemma in Section 4.1.1.

---

coincides with our formulation.

Lemma 3 (Replacement Lemma for Randomized Constant Price Paths). *Suppose that the seller randomizes over constant price paths, while nature provides information independently of the realized price. Then the seller's profit can be minimized by an information structure that only provides information in period* 1.

This result embeds Lemma 1 as a special case when the seller charges a deterministic constant price. However, the "push and discount" argument we used there to find a replacement static information structure does not readily extend to the current setting. This is because with random prices, nature's recommendation in a given period is not just a binary decision to purchase or not; rather, any signal suggests a set of prices at which the buyer should purchase. Such information is higher-dimensional than in the deterministic case, and we need new tools to generalize the previous argument.

To address this difficulty, for any given dynamic information structure, we seek a replacement static information structure that induces the same discounted probability of sale *conditional on each possible price realization.* In our proof of Lemma 3, we introduce the concept of "cutoff prices" for a given price-independent information process. These cutoff prices are the dual notion of "cutoff values" used in the proof of Lemma 2: They represent the highest prices at which the consumer would purchase in period t, given the information up until that time and the expected future information.

It turns out that the distribution over cutoff prices is sufficient to determine the probability of sale given any constant price path. Specifically, analogous to (3), the seller's total profit (conditional on any realized price) can be written as a discounted sum of one-period profits from buyers whose values are given by the cutoff prices. Therefore, the same profit is obtained if the buyer is simply informed of the cutoff price at a random period, drawn according to a Geometric(δ) distribution.[31] The remaining challenge is to show that this distribution of cutoff prices can be induced as the buyer's posterior expected values under some static information structure. We prove this by applying the mean-preserving spread characterization of Rothschild-Stiglitz (1970), with some additional technical details that we explain in Online Appendix B.

6. PRICE DYNAMICS UNDER RESTRICTIONS ON INFORMATION PROCESSES

Our main result provides a clear prescription for a monopolist who is completely uncertain about how consumers will learn about his product: Keep the price fixed over time at the single-period optimum. In this section, we consider two modifications of our main model, where our reduction

---

[31]When the seller charges a deterministic constant price p, the cutoff price first exceeds p precisely in the period when the buyer would purchase under the original dynamic information structure. Thus in that special case, the current proof reduces to the "push and discount" argument in Section 4.1.1.

to the one-period problem fails and dynamic selling strategies become optimal in the presence of learning. In Section 6.1, we show how declining prices out-perform constant prices when the seller believes that learning does not occur in every period (e.g., when information is somewhat rare). Section 6.2 shows that introductory pricing is favored when buyers with common values arrive over time and information is publicly observed.

6.1. Infrequent Information

In our main model information can arrive in each period. Here we study a stylized variant where $T = 2$ and information is *constrained to only arrive in one of the two periods*. Formally, we restrict to dynamic information structures (as defined in Section 2.1) with either signal set S_1 or S_2 being a singleton. This captures a setting where information is infrequent, and learning may not occur every period. The following result shows that the seller can now obtain a profit guarantee higher than Π^* with a decreasing price path. As a corollary, the optimal deterministic pricing strategy involves decreasing prices. This result highlights that the optimality of constant price paths relies on the seller seeking robustness against *dynamic* information structures.

Proposition 4. *Suppose that $T = 2$ and that the buyer either receives information in period one or period two, but not both. Further suppose $p^* > \underline{v}$. Then for any $\delta \in (0,1)$, there exists a price path $p_1 > p_2 = p^*$ that guarantees profit strictly greater than Π^*.*

The intuition for this proposition goes back to the upper-bound argument (Lemma 2) in Section 4.1.2. There we showed how nature could use a threshold information process to hold profit below Π^*. Against a decreasing price path, the constructed process involved two thresholds, one in each period. However, only one threshold is allowed in the current setting. If nature were to remove the threshold in the first period, then the buyer would purchase at the slightly higher price p_1 to avoid the cost of discounting.[32] But if nature were to remove the threshold in the second period, then the probability of sale would jump up in that period unless $p^* = \underline{v}$. Either way, profit would strictly exceed Π^*, suggesting that nature can only hold the seller to the single-period profit level by utilizing dynamic information.

---

[32]Note that given any price, the worst-case static information structure induces the same amount of buyer surplus as no information. So in this problem, when p_1 is equal to p_2, the buyer strictly prefers to purchase in period one (without any information) than to purchase later (facing worst-case information). By continuity, the same holds for p_1 slightly larger than p_2.

6.2. Common Values and Public Information

This subsection shows that informational interdependence across buyers can favor increasing prices.[33] To discuss this possibility, we consider here multiple buyers who arrive over time. Note that arriving buyers by itself presents no change to our constant price path result. Indeed, constant prices guarantee the one-period maxmin profit from each buyer, delivering a lower-bound on the seller's total profit. On the other hand, so long as buyer values are independent or private information structures are allowed, nature can minimize the seller's profit from each buyer simultaneously. Under either of these assumptions, constant price paths would remain optimal.

This argument (in particular, the profit upper-bound) is no longer valid if buyers share both value and information. Below we assume that all arriving buyers have the same value for the product, and that all information is public to the buyers. More formally, we consider the following interaction:

- First, the seller chooses a pricing strategy $\sigma \in \Delta(P^T)$

- Next, nature chooses an information process $\mathcal{I} = (I_t)_{t=1}^T$, with $I_t : V \times S^{t-1} \times P^t \to \Delta(S_t)$.

- The value for the object is drawn, with $v \sim F$.

- One new buyer arrives in each period $t = 1, 2, \ldots, T$. All buyers value the object at v.

- Upon arrival (and in every period until they purchase), each buyer observes $p^t = (p_1, \ldots, p_t)$ and $s^t = (s_1, \ldots, s_t)$. In every period, any buyer who has arrived and not purchased can either purchase or delay, with payoffs discounted by a factor δ.

- The seller chooses the pricing strategy assuming that $\mathcal{I}$ minimizes total discounted profit.

The key distinction from our main model is that nature is more restricted when minimizing the seller's profit. Choosing an information structure for one buyer will influence the profit obtained from later buyers, who will observe the entire signal history.

We characterize the seller's profit guarantee per buyer in the patient limit, which establishes an interesting connection to the results in Section 5.

Proposition 5. *Consider the model with common values and public signals. Let $\Pi^C(\delta, T)$ be the seller's maxmin discounted total profit with discount factor δ and time horizon T. We have:*

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

---

[33]Optimal pricing when information is conveyed across buyers has been studied using the Bayesian approach, such as in Bose et al. (2006, 2008). A key distinction is that we allow buyers to delay purchase.
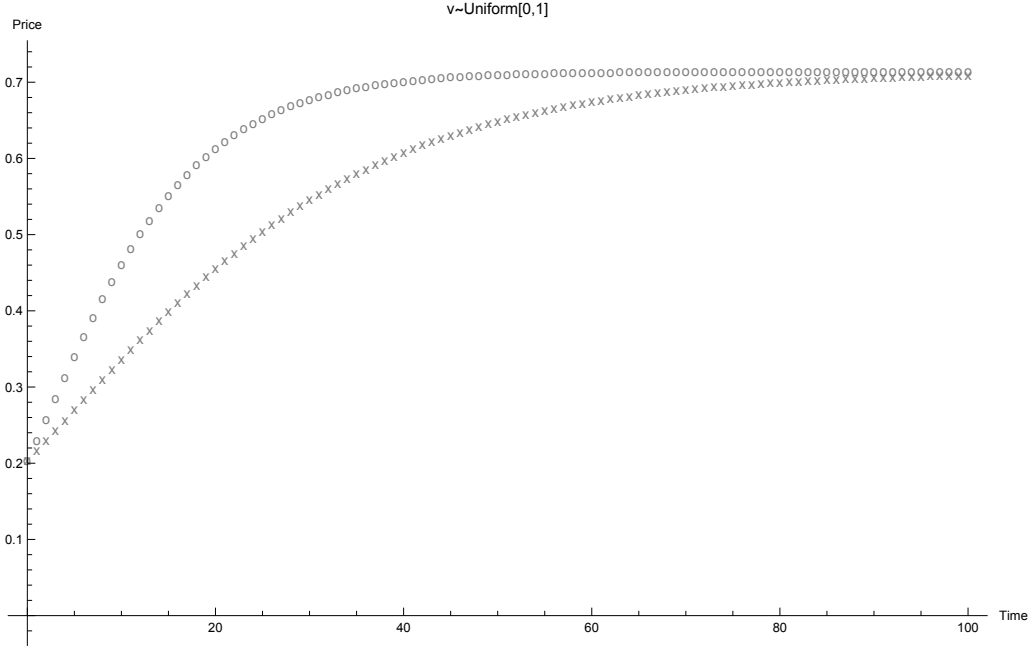
Figure 1: Illustration of price paths. The "o" path is $\delta = 0.9$; the "x" path is $\delta = 0.95$.

This profit can be approximated by a sequence of strictly increasing price paths.

Figure 1 illustrates the price paths we use for this approximation, in the case of a uniform prior. For each δ, prices start off at Π_{RSD}, increase over time, and converge as $T \to \infty$ to the number S from (6), where $S \approx 0.715$ for the uniform distribution.[34]

To see why Proposition 5 holds, we first observe that nature can provide the Roesler-Szentes information structure in the first period and hold profit below Π_{RSD} per buyer. In the opposite direction, we look for increasing price paths that guarantee close to Π_{RSD}. The following analogue of the Replacement Lemma greatly simplifies the analysis:

Lemma 4 (Replacement Lemma for Common Values). *Consider the model with common values and public signals. Suppose that the seller uses a deterministic and increasing price path. Then total profit can be minimized by an information structure that only provides information in period 1.*

Lemma 4 enables us to restrict attention to static information structures. To complete the proof, we adapt Du's random price distribution (6) to construct (deterministic) price paths for which the profit under any static information structure approximates the single-period profit under Du's mechanism. As a consequence, per buyer profit guarantee converges to Π_{RSD}.

---

[34]To compute this, we use equation (24) in Online Appendix B. There we show that $\int_0^{\hat{S}} F_W^B(v)\, dv \leq \int_0^{\hat{S}} F(v)\, dv$, with equality at $\hat{S} = S$. The first order condition implies $F_W^B(S) = F(S)$, which gives $1 - \frac{W}{S} = S$ in the uniform case. Since $W \approx 0.2037$ as pointed out in Roesler and Szentes (2017), we deduce $S \approx 0.715$.

7. CONCLUSION

In this paper, we have utilized a robust approach to study optimal monopoly pricing with dynamic information arrival. In our baseline model, the monopolist's optimal profit guarantee is what he would obtain with only a single period to sell, and a constant price path delivers this optimal profit. The lesson from our paper is thus that, when seeking robustness against a sufficiently rich class of information arrival processes, the dynamic problem reduces to the static one, as in the known-values case. This provides a useful benchmark, since performing a Bayesian analysis with general information structures would typically disallow a parsimonious result similar to our Theorem 1 (see the example in Section 1.1). We also identify several economically meaningful restrictions on the informational environment that would lead to gains from non-constant pricing strategies.

Our baseline model describes settings where the seller shares the buyer's prior about her value, but does not know how her expected value will evolve over time. For the car buyer discussed in the Introduction, this reflects that both seller and buyer understand the overall distribution of breakdown probabilities, but the seller faces uncertainty about what the mechanic will convey about the buyer's idiosyncratic situation. In some other applications, it may be a strong assumption that the value distribution is common knowledge but information is not. To address this concern, we have discussed how our analysis extends to sellers facing distributional uncertainty on one hand (Online Appendix D), as well as to sellers possessing some knowledge of the buyer's information on the other (Section 4.2). The latter extension demonstrates how our constant price path result can be seen as a strict generalization of the known-values setting. We hope this connection between the Bayesian and robust modeling approaches will be further explored in future work.

We view one contribution of this paper as introducing a robust objective into a dynamic mechanism design problem. Dynamics complicate the characterization of agent behavior, which is essential for understanding the performance of a given mechanism across different (informational) environments. 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. But in terms of economic motivation, dynamic robustness concerns are also present in other applications. The techniques developed in this paper may help other researchers further extend the robust mechanism design literature to accommodate dynamics.

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. Extending by continuity, we define $g(0) = \underline{v}$ to be the (essential) minimum of the value distribution F.*

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

We note that if F does not have a mass point at $\underline{v}$, then $g(\alpha)$ 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 $\beta < \underline{v}$, after which it jumps to m and increases continuously to 1. Thus even when F is discrete, the pressed distribution G is continuous except possibly at $\underline{v}$.

While intuitive, the "lowest α-percentile" of the distribution F can be formally defined as follows. If there exists w such that $F(w) = \alpha$, then v is in the lowest α-percentile if and only if $v \leq w$. Otherwise, there must exist w such that $F(w_-) < \alpha < F(w)$. Let $m = F(w) - F(w_-)$ be the mass at w, and consider a random variable U_w that is independent of the true value v and uniformly distributed on $[0, m]$. Then the lowest α-percentile are those pairs (v, U_w) such that $v < w$, or $v = w$ and $U_w \leq \alpha - F(w_-)$.

Next, we also present a generalized definition of threshold processes:

DEFINITION 2'. *A (descending) threshold information process involves a descending sequence of (possibly randomized) percentiles $1 \geq \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_T \geq 0$, where each α_t is measurable with respect to realized prices $p_1, \ldots, p_t$. Under this process, in each period t the buyer is told whether or not v is in the lowest α_t-percentile of the distribution F.*

In case the prior distribution F has atoms, this definition allows nature to "break the atom" in providing information. For example, if F is supported on two values $\underline{v} < \overline{v}$, then nature could tell the buyer whether 1) she is among a particular half of the buyers with value $\overline{v}$, or 2) she is either among the other half of the buyers with value $\overline{v}$ or among those with value $\underline{v}$.

The rest of this appendix provides proofs for Proposition 1, Proposition 2, Theorem 1, Proposition 3 and Theorem 1'.

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 $\hat{\mathcal{I}}$, 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 purchase or not. Nature chooses an information structure that minimizes the probability of "purchase." By Lemma 1 in Kolotilin (2015), this minimum is achieved by a threshold information structure, namely by recommending purchase for $v > y$ and not for $v \leq y$. Since the buyer's expected value given $v \leq y$ cannot be greater than p, we have $y \leq F^{-1}(G(p))$. It is then easy to see that the particular information structure $\mathcal{I}$ above, which sets $y = F^{-1}(G(p))$, 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. From earlier discussion, we know that G is (almost) continuous. Hence $p^* = \arg\max_p p(1 - G(p))$ exists except when F has a mass point at $\underline{v}$ and $\underline{v} > p(1 - G(p)), \forall p$. In the latter case (for example when F is a point-mass), the maxmin profit of $\underline{v}$ is not achievable due to tie-breaking. But for any $\epsilon > 0$, the seller can guarantee profit $\underline{v} - \epsilon$ by choosing $p = \underline{v} - \epsilon$. Our subsequent results about the dynamic model continue to hold, so long as the seller's $\max\min$ objective is replaced with $\sup\inf$.

A.2. Proof of Proposition 2

The profit comparison $\Pi^* \leq \hat{\Pi}$ is straightforward, because nature can always provide full information to the buyer, so that

$$\Pi^* = p^*(1 - G(p^*)) \leq p^*(1 - F(p^*)) \leq \hat{\Pi}.$$

Equality requires $p^* = \underline{v}$ (otherwise $G(p^*)$ is strictly bigger than $F(p^*)$), as well as $\hat{p} = p^*$ (otherwise the second inequality is strict).

The price comparison $p^* \leq \hat{p}$ is more difficult to show. We first present the proof assuming that the distribution F is continuous. 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,

$$pG(p) = \int_0^{F^{-1}(G(p))} v \, dF(v), \forall p \in [\underline{v}, \mathbb{E}[v]]. \tag{7}$$

Differentiating both sides with respect to p, we obtain

$$G(p) + pG'(p) = \frac{\partial}{\partial p}(F^{-1}(G(p))) \cdot F^{-1}(G(p)) \cdot F'(F^{-1}(G(p))) = G'(p) \cdot F^{-1}(G(p)). \quad (8)$$

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

$$G'(p) = \frac{G(p)}{F^{-1}(G(p)) - p}. \quad (9)$$

Thus,

$$G(p) + pG'(p) = \frac{G(p) \cdot F^{-1}(G(p))}{F^{-1}(G(p)) - p}. \quad (10)$$

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

$$\frac{p}{1 - G(p)} \cdot \left(1 - F\left(\frac{p}{1 - G(p)}\right)\right) < p. \quad (11)$$

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.

For a general (potentially discrete) distribution F, the pressed distribution G is not necessarily differentiable, and we need to proceed more carefully. Given a price p, define

$$x(p) = \min\{v : \ F(v) \geq G(p)\}.$$

The minimum exists because the c.d.f. F is right-continuous. This $x(p)$ will play the role of $F^{-1}(G(p))$ in the above analysis.

Specifically, we now show that similar to (9), the *left-derivative* of G at p is given by $\frac{G(p)}{x(p)-p}$. Formally, consider any small positive number ϵ. Recall that $p \cdot G(p)$ is the integral of values in the lowest $G(p)$-percentile of the distribution F, and $(p - \epsilon) \cdot G(p - \epsilon)$ is the corresponding integral in the lowest $G(p - \epsilon)$-percentile. Thus the difference $p \cdot G(p) - (p - \epsilon) \cdot G(p - \epsilon)$ is the integral of values between the $G(p - \epsilon)$- and $G(p)$-percentile. By the definition of $x(p)$, the values between

these two percentiles are close to $x(p)$ as $\epsilon \to 0$.[35] We can thus write

$$p \cdot G(p) - (p - \epsilon) \cdot G(p - \epsilon) = (G(p) - G(p - \epsilon)) \cdot (x(p) + o(1)),$$

where $o(1)$ is a vanishing term as $\epsilon \to 0$. Rearranging, we obtain

$$\epsilon \cdot G(p - \epsilon) = (x(p) + o(1) - p) \cdot (G(p) - G(p - \epsilon))$$

It follows that
$$\frac{G(p) - G(p - \epsilon)}{\epsilon} = \frac{G(p - \epsilon)}{x(p) + o(1) - p} \to \frac{G(p)}{x(p) - p},$$

as we desire to show.

Since $G(p)$ is left-differentiable, so is the profit function $p(1 - G(p))$, whose left-derivative is computed to be (similar to (10))

$$1 - G(p) - pG'_{\text{left}}(p) = 1 - \frac{G(p) \cdot x(p)}{x(p) - p}.$$

If we can show $x(p) < \frac{p}{1 - G(p)}$ for $p > \hat{p}$, then $G(p)x(p) > x(p) - p$ and so the profit function has negative left-derivative. This will be sufficient to imply that $p(1 - G(p))$ is strictly decreasing for $p \geq \hat{p}$.[36] Hence $p^* \leq \hat{p}$.

Recall that $x(p)$ is defined to be the smallest v such that $F(v) \geq G(p)$. So in order to show $x(p) < \frac{p}{1 - G(p)}$, we only need to show

$$F\left(\left(\frac{p}{1 - G(p)}\right)_-\right) > G(p),$$

where the LHS represents $\lim_{\epsilon \to 0} F(\frac{p}{1 - G(p)} - \epsilon)$. Similar to what we did in the continuous distribution case, the above inequality can be rewritten as

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

---

[35]From the definition, for any $\delta > 0$ it holds that $F(x(p) - \delta) < G(p)$. Thus for ϵ small, $F(x(p) - \delta) < G(p) - \epsilon$, which implies that the value at the $G(p - \epsilon)$-percentile is at least $x(p) - \delta$. On the other hand, since $F(x(p)) \geq G(p)$, the value at the $G(p)$-percentile is at most $x(p)$.

[36]Indeed, it suffices to show that the function $\Pi(p) = p(1 - G(p))$ is injective and thus strictly monotone for $p \geq \hat{p}$. This can be proved similar to Rolle's Theorem: Suppose for contradiction that $\Pi(p_1) = \Pi(p_2)$ at some prices $p_2 > p_1 \geq \hat{p}$. Since the left-derivative at p_2 is negative, $\Pi(p_2 - \epsilon) > \Pi(p_2)$ for ϵ small. Thus the continuous function Π has an interior maximizer on the interval $[p_1, p_2]$. But then the left-derivative at this maximizer must be non-negative, leading to a contradiction.

This holds because the LHS is the profit from charging $\frac{p}{1-G(p)}$ under known values. By definition of $\hat{p}$, the profit is indeed bounded above by $\hat{p}(1 - F(\hat{p}_-)) \le \hat{p} < p$. Finally, whenever $\hat{p} > \underline{v}$ we have $F(\hat{p}_-) > 0$, and so the above strict inequality holds even at $p = \hat{p}$. Therefore $p(1 - G(p))$ has negative left-derivative at $p = \hat{p}$, so that p^* is strictly smaller than $\hat{p}$. This completes the proof.

A.3. Proof of Theorem 1

As discussed in the main text, the proof consists of a lower-bound and an upper-bound. For the lower-bound on the seller's profit guarantee, we will prove Lemma 1. This is sufficient to imply that when the seller charges a constant price path of p^*, his profit is minimized by a static information structure which induces no delay. Thus by our one-period analysis, the seller can guarantee Π^*. As for the upper-bound, we will prove Lemma 2, which directly constructs an information process (for any pricing strategy) that holds profit below Π^*. We address these two parts in turn.

A.3.1. Lower-bound: Proof of Lemma 1

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 previous signals (and not on prices). We can also think about the stopping time τ as a function of signal realizations.

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 $S = \{s_H, s_L\}$, corresponding to the recommendation to purchase or not, respectively. To specify the distribution of these signals conditional on the true value 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 purchasing the object, let the buyer receive the signal s_H with probability $\delta^{\tau-1}$. With complementary probability and when $\tau = \infty$, let her receive the other signal s_L. In the alternative information structure $\mathcal{I}'$, nature reveals s_H or s_L in the first period and provides no more information afterwards.

We claim that under $\mathcal{I}'$, the buyer receiving the signal s_L has expected value at most p_1. To this end, define $y_t = \mathbb{E}[v \mid \tau = t]$ be the buyer's expected value when she is recommended to purchase in period t, under the original information structure. This definition applies to $1 \le t \le T$ as well as $t = \infty$, in which case y_∞ is the buyer's expected value in case she is never recommended to purchase. Note that under the original information structure, stopping at time τ must be weakly better than stopping at time 1. Thus

$$\mathbb{E}[v] - p_1 \le \mathbb{E}\left[\delta^{\tau-1} \cdot (y_\tau - p_\tau)\right], \tag{12}$$

where the RHS expectation is taken with respect to the distribution of the stopping time τ.

Since $p_\tau \geq p_1$, simple algebra reduces (12) to the following:

$$\mathbb{E}[v] \leq \mathbb{E}\left[\delta^{\tau-1}y_\tau + (1 - \delta^{\tau-1})p_1\right]. \tag{13}$$

Observe that $\mathbb{E}[v] = \mathbb{E}\left[\mathbb{E}[v \mid \tau]\right] = \mathbb{E}[y_\tau]$. Thus the above inequality implies $\mathbb{E}\left[(1 - \delta^{\tau-1}) \cdot y_\tau\right] \leq \mathbb{E}[(1 - \delta^{\tau-1})p_1]$. That is,

$$p_1 \geq \frac{\mathbb{E}\left[(1 - \delta^{\tau-1}) \cdot y_\tau\right]}{\mathbb{E}[1 - \delta^{\tau-1}]}. \tag{14}$$

The denominator $\mathbb{E}[1 - \delta^{\tau-1}]$ can be rewritten as $\mathbb{P}[s_L]$, which is the probability of receiving s_L under the replacement information structure $\mathcal{I}'$. On the other hand, the numerator in (14) equals

$$\mathbb{E}\left[\mathbb{P}[s_L \mid \tau] \cdot \mathbb{E}[v \mid \tau]\right],$$

which can be further rewritten as

$$\mathbb{E}\left[\mathbb{P}[s_L \mid \tau] \cdot \mathbb{E}[v \mid s_L, \tau]\right],$$

because s_L does not provide more information about v beyond τ.

With these, the above inequality (14) states that

$$p_1 \geq \frac{\mathbb{E}\left[\mathbb{P}[s_L \mid \tau] \cdot \mathbb{E}[v \mid s_L, \tau]\right]}{\mathbb{P}[s_L]} = \mathbb{E}[v \mid s_L] \tag{15}$$

just as we claimed.

Thus, under the static information structure $\mathcal{I}'$ constructed above, a buyer who receives the signal s_L has expected value at most p_1, which is also less than any future price. This buyer does not purchase under $\mathcal{I}'$. Furthermore, a buyer observes s_H only if purchasing was incentive compatible under the original information structure; since $\mathbb{E}[v \mid \tau = t] \geq p_t \geq p_1$ for all t, we have $\mathbb{E}[v \mid s^H] \geq p_1$. Such a buyer purchases in the first period under $\mathcal{I}'$ (as there is no future information). It follows that the probability of sale under the replacement information structure is $\mathbb{E}[\delta^{\tau-1}]$, and the seller's profit is $\mathbb{E}[\delta^{\tau-1}] \cdot p_1$. This is no more than $\mathbb{E}[\delta^{\tau-1} \cdot p_\tau]$, the discounted profit under the original dynamic information structure. Hence the lemma.

A.3.2. Upper-bound: Proof of Lemma 2

In the main text we sketched an argument to prove Lemma 2 for deterministic price paths. Here we provide a formal treatment of the general case, where the pricing strategy σ may be randomized.

For clarity, the proof is broken down into four steps.

Step 1: Cutoff values. 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

$$v - p_t \geq \max_{\tau \geq t+1} \mathbb{E}\left[\delta^{\tau-t} \cdot (v - p_\tau) \mid p_1, \ldots, p_t\right], \tag{16}$$

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 $v \geq v_t$.[37] Thus, v_t is defined by the equation

$$v_t - p_t = \max_{\tau \geq t+1} \mathbb{E}\left[\delta^{\tau-t} \cdot (v_t - p_\tau) \mid p_1, \ldots, p_t\right], \tag{17}$$

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 $t \geq 1$

$$w_t = \min\{v_1, v_2, \ldots, v_t\} = \min\{w_{t-1}, v_t\}. \tag{18}$$

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

Step 2: Construction of information process. Consider the following threshold information process $\mathcal{I}$. 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 pricing strategy.

Step 3: Buyer behavior under this process. The following lemma describes the buyer's optimal stopping decision in response to σ and $\mathcal{I}$:

Lemma 5 (Optimal Stopping). *For any pricing strategy σ, let the information process $\mathcal{I}$ be constructed as above. Then the buyer finds it optimal to follow nature's recommendation: She purchases in the first period when told her value is above the $G(w_t)$-percentile (and waits otherwise).*

To prove this lemma, 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 (18). Given this

---

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

signal, the buyer 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 purchases 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 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. Lemma 5 follows.

Step 4: Profit decomposition. By Lemma 5, the buyer whose true value belongs to the percentile range $(G(w_t), G(w_{t-1})]$ will purchase 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 6 (Price Equals Discounted Cutoffs). *Suppose $w_t = v_t \leq w_{t-1}$ in some period t. Then*

$$p_t = \mathbb{E}\left[\sum_{s=t}^{T-1} (1-\delta)\delta^{s-t} w_s + \delta^{T-t} w_T \mid p^t\right] \tag{19}$$

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

Using Lemma 6, we can rewrite the profit as

$$
\begin{aligned}
\Pi &= \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} \cdot (G(w_{t-1}) - G(w_t)) \cdot \mathbb{E}\left[\sum_{s=t}^{T-1} (1-\delta)\delta^{s-t} w_s + \delta^{T-t} w_T \mid p^t\right]\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} \cdot (G(w_{t-1}) - G(w_t)) \cdot \left(\sum_{s=t}^{T-1} (1-\delta)\delta^{s-t} w_s + \delta^{T-t} w_T\right)\right] \\
&= \mathbb{E}\left[\sum_{s=1}^{T-1} (1-\delta)\delta^{s-1} w_s(1 - G(w_s)) + \delta^{T-1} w_T(1 - G(w_T))\right] \\
&\leq \Pi^*.
\end{aligned}
\tag{20}
$$

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 the upper-bound, it only remains to show Lemma 6.

Proof of Lemma 6. We assume that T is finite,[38] and prove the result by induction on $T - t$. The base case $t = T$ follows from $w_T = v_T = p_T$. For $t < T$, from (17) we can find an optimal stopping time $\tau \geq t + 1$ such that[39]

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

which can be rewritten as

$$p_t = \mathbb{E}[(1 - \delta^{\tau-t})v_t + \delta^{\tau-t}p_\tau \mid p^t]. \tag{21}$$

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 (18); while in period τ (if $\tau < \infty$), $v_\tau \leq v_t$ 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

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

And if $\tau \leq T$, we can 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 equations into (21) proves Lemma 6 as well as Theorem 1. □

A.3.3. Example: Profit Can be Even Worse

The threshold information process in the upper-bound argument directly generalizes the one-period construction. Despite this analogy, however, this particular process is generally not the worst case beyond a single period. Here we provide a concrete example to illustrate:

---

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

[39]For completeness, we define $p_\infty = 0$.

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 threshold information process constructed before Lemma 5 induces expected value $\frac{9}{20}$ when recommending the buyer not to purchase in the first period. This information process further induces expected value $p_2 = 1/10$ when recommending the buyer not to purchase in the second period.*

If the probability of being recommended to purchase 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 $\frac{9}{20} = r_2 + \frac{1}{10}(1 - r_2)$ because beliefs are martingales. Thus we obtain $r_1 = \frac{1}{11}$ and $r_2 = \frac{7}{18}$. Profit under this information process 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

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

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

The intuitive explanation for this example is that nature can promise more information (relative to our constructed process) to the buyer in the second period. This creates option value and induces delay, which hurts the seller's profit when price in the second period is much lower. In light of Lemma 1, prices declining over time are crucial for such an example. Conversely, this example also shows that the Replacement Lemma only holds with non-decreasing prices.

A.3.4. Unique Optimality with Arriving Buyers

As Lemma 1 shows, the seller can guarantee profit at least Π^* from a single buyer using any increasing price path that starts with p^*. However, a strategy involving strictly increasing prices would not be optimal in case additional buyers were to arrive after period 1. To make this point most clear, we consider here a situation in which *one buyer arrives in each period*, with value independently drawn from the prior distribution F. This buyer then learns about her value over time according to some information process, and optimally decides when to purchase.

Note that the independent values assumption distinguishes from the model considered in Section 6.2. Specifically, in the current model nature can release information to minimize the profit from different buyers simultaneously. It follows that the seller's total profit guarantee cannot

exceed $\sum_{t=1}^{T} \Pi^* \cdot \delta^{t-1} = \Pi^* \cdot \frac{1-\delta^T}{1-\delta}$. On the other hand, a constant price path of p^* makes the environment stationary and achieves this total profit guarantee.

The following result additionally shows that constant pricing is the *uniquely optimal* strategy in this setting:

Proposition 6. *Suppose $p^* = \arg\min_p p(1 - G(p))$ is unique in the one-period problem. Then in the model with one buyer arriving in each period (with independent values), the constant price path of p^* uniquely achieves the maxmin total profit $\Pi^* \cdot \frac{1-\delta^T}{1-\delta}$.*

Proof. Since nature can independently minimize the profit from different buyers, any pricing strategy that guarantees total profit $\Pi^* \cdot \frac{1-\delta^T}{1-\delta}$ must guarantee $\Pi^* \cdot \delta^{t-1}$ from the buyer arriving in period t, for each $1 \leq t \leq T$. In particular, profit from the first buyer must equal Π^*. From inequality (20) above, we see this can only occur if $w_s = p^*$ almost surely for each s. Thus by Lemma 6, $p_1 = p^*$ with probability one. Similar consideration for later buyers shows that the seller must always charge p^* to achieve the total profit guarantee $\Pi^* \cdot \frac{1-\delta^T}{1-\delta}$. $\qquad\square$

We note that the assumption of p^* being unique in the one-period problem is satisfied for *generic* distributions F. Alternatively, uniqueness is guaranteed when the function $p(1 - G(p))$ is strictly quasi-concave in p, which can be ensured by a regularity condition on F (see Lemma 8 in Online Appendix D).

A.4. Proof of Proposition 3

In the proof below, we fix an arbitrary information process, and then construct a threshold process that leads to lower profit. For ease of exposition, we first assume the prior distribution F is continuous.

Step 1: Construction of the threshold process. To begin, we assume the original (price-dependent) dynamic information structure simply recommends the buyer to purchase or not conditional on the realized prices so far. This simplification is known as the "revelation principle for information design"; see Makris and Renou (2019) for a general treatment. Given such an information process, we define for $1 \leq t \leq T$ a random variable λ_t, which is the probability that the buyer is recommended to purchase in period t (λ_t is adapted to the realized prices p^t). Also define y_t to be the buyer's expected value given this recommendation. We further define λ_{T+1} to be the probability that the buyer is never recommended to purchase, and y_{T+1} to be the expected value conditional on this event.

We then construct a threshold process with price-dependent thresholds

$$\infty = v_0 \geq v_1 \geq v_2 \geq \cdots \geq v_T \geq v_{T+1} = 0,$$

such that v_t depends (only) on realized prices p^t, and $\mathbb{P}[v_t < v \le v_{t-1} \mid p^t] = \lambda_t$ for every t and every price history p^t. When the prior distribution F is continuous, the condition $F(v_{t-1}) - F(v_t) = \lambda_t$ can be used to choose the thresholds v_t iteratively, starting from $t = T + 1$ to smaller t.

Let z_t denote the average value conditional on v belonging to the interval $(v_t, \ v_{t-1}]$, then we have the following inequality holding for every price history p^T:

$$\sum_{r=t+1}^{T+1} \lambda_r \cdot y_r \ge \sum_{r=t+1}^{T+1} \lambda_r \cdot z_r, \quad \forall 0 \le t \le T. \tag{22}$$

This follows from a key property of threshold information structures: Given a mass $\sum_{r=t+1}^{T+1} \lambda_r$ of buyers, their average value is minimized when they are precisely those buyers with value less than v_t, and this average is the RHS of (22) divided by $\sum_{r=t+1}^{T+1} \lambda_r$.

Step 2: Buyer incentives. Using (22), we are going to show that when the buyer learns her value is *below* v_t, she optimally delays purchase. To see this, consider a buyer who is recommended not to purchase in period t under the *original* process $\mathcal{I}$. Incentive compatibility requires

$$\mathbb{E}\left[\sum_{s=t+1}^{T+1} \lambda_s \cdot (y_s - p_t) \mid p^t \right] \le \mathbb{E}\left[\sum_{s=t+1}^{T} \delta^{s-t} \lambda_s \cdot (y_s - p_s) \mid p^t \right].$$

Rearranging, this yields

$$\mathbb{E}\left[\sum_{s=t+1}^{T} (1 - \delta^{s-t}) \lambda_s y_s + \lambda_{T+1} y_{T+1} \mid p^t \right] \le \mathbb{E}\left[\sum_{s=t+1}^{T+1} \lambda_s p_t - \sum_{s=t+1}^{T} \delta^{s-t} \lambda_s p_s \mid p^t \right].$$

Observe that $\sum_{s=t+1}^{T} (1 - \delta^{s-t}) \lambda_s y_s + \lambda_{T+1} y_{T+1}$ on the LHS above can be written as a positive linear combination of the LHS of (22) for different t. So we can use (22) to replace y_s by z_s everywhere without changing the inequality:

$$\mathbb{E}\left[\sum_{s=t+1}^{T} (1 - \delta^{s-t}) \lambda_s z_s + \lambda_{T+1} z_{T+1} \mid p^t \right] \le \mathbb{E}\left[\sum_{s=t+1}^{T+1} \lambda_s p_t - \sum_{s=t+1}^{T} \delta^{s-t} \lambda_s p_s \mid p^t \right].$$

Rearranging again gives

$$\mathbb{E}\left[\sum_{s=t+1}^{T+1} \lambda_s \cdot (z_s - p_t) \mid p^t \right] \le \mathbb{E}\left[\sum_{s=t+1}^{T} \delta^{s-t} \lambda_s \cdot (z_s - p_s) \mid p^t \right].$$

This shows that under the threshold process, a buyer with value below v_t should not purchase in period t.

Step 3: Profit comparison. By the above analysis, the threshold process $\mathcal{I}'$ ensures that any buyer with value in $(v_t, v_{t-1}]$ purchases in period t or later. If she indeed purchased in period t, expected discounted profit from such a buyer would be $\delta^{t-1}\lambda_t p_t$ (conditional on p^t). If this buyer were to delay purchase until later, discounted profit would be even lower because social surplus decreases (due to discounting) while buyer surplus could only increase.

Hence we have shown that the seller's total expected profit under the threshold information process is bounded above by $\mathbb{E}[\sum_{t=1}^{T} \delta^{t-1}\lambda_t p_t]$, which is what he obtains under the original information structure. This proves Proposition 3 when F is a continuous distribution.

To prove the result for a general distribution F, we recall Definition 2' at the beginning of Appendix A, which provides a suitably generalized definition of threshold processes. We can carry out essentially the same proof as above, except that in the general case we let the threshold process inform the buyer whether or not her value is in the lowest $\sum_{r=t+1}^{T+1} \lambda_r$-percentile of the distribution F. The rest of the proof holds without change.

A.5. Proof of Theorem 1'

On one hand, the Replacement Lemma implies that when using a constant price path of p, the seller obtains profit at least $p(1 - G_{\underline{s}}(p))$ from the buyer with initial signal realization $\underline{s}$. Thus the seller's expected profit across different initial signals is at least $p(1 - \mathbb{E}[G_{\underline{s}}(p)])$, which yields the profit guarantee $\Pi_{\mathcal{H}}^*$ when optimally choosing $p = p_{\mathcal{H}}^*$.

On the other hand, we can generalize Lemma 2 to show that for any pricing strategy, there exists a collection of information processes (one for each initial signal realization) that hold expected profit below $\Pi_{\mathcal{H}}^*$. In fact, we can follow the same steps as in the proof of Lemma 2: In each period t, a buyer with initial signal $\underline{s}$ is told whether or not her value is in the lowest $G_{\underline{s}}(w_t)$-percentile of the distribution $F_{\underline{s}}$. Such a buyer purchases if and only if her value is above this percentile. Note that the binding cutoff values w_t as defined in (17) and (18) depend on the pricing strategy, but do not depend on the initial signal realization.

Under this information structure, total profit can be computed (similar to before) as

$$\Pi = \mathbb{E}_{\underline{s}}\left[\mathbb{E}_{p,w}\left[\sum_{t=1}^{T} \delta^{t-1} \cdot (G_{\underline{s}}(w_{t-1}) - G_{\underline{s}}(w_t)) \cdot p_t \mid \underline{s}\right]\right].$$

To be precise, $\mathbb{E}_{p,w}$ in the above equation indicates that the inner conditional expectation is taken with respect to the process of prices and cutoffs, which are independent of $\underline{s}$ (but $\underline{s}$ does affect

the pressed distribution $G_{\underline{s}}$). The outer expectation $\mathbb{E}_{\underline{s}}$ then computes the average profit across different possible initial signals $\underline{s}$.

We can use the equalities in (20) to simplify the inner expectation above, and then exchange the order of taking expectations. This yields

$$\Pi = \mathbb{E}_{\underline{s}}\left[\mathbb{E}_w\left[\sum_{t=1}^{T-1}(1-\delta)\delta^{t-1}w_t(1-G_{\underline{s}}(w_t)) + \delta^{T-1}w_T(1-G_{\underline{s}}(w_T)) \mid \underline{s}\right]\right]$$
$$= \mathbb{E}_w\left[\sum_{t=1}^{T-1}(1-\delta)\delta^{t-1}w_t(1-\mathbb{E}_{\underline{s}}[G_{\underline{s}}(w_t)]) + \delta^{T-1}w_T(1-\mathbb{E}_{\underline{s}}[G_{\underline{s}}(w_T)])\right].$$

Note that we use "t" to replace the role of "s" in the penultimate line of (20), in order to avoid confusion with the initial signal "$\underline{s}$."

Since by definition $w_t(1-\mathbb{E}_{\underline{s}}[G_{\underline{s}}(w_t)]) \leq \Pi_{\mathcal{H}}^*$ holds for each w_t, the last displayed equation implies $\Pi \leq \Pi_{\mathcal{H}}^*$, as we desire to show.

References

Ricardo Alonso and Odilon Câmara. Political disagreement and information in elections. *Games and Economic Behavior*, 100: 390–412, 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.

Dirk Bergemann and Achim Wambach. Sequential information disclosure in auctions. *Journal of Economic Theory*, 159: 1074–1095, 2015.

David Blackwell. Equivalent comparison of experiments. *The Annals of Mathematical Statistics*, 24(2): 265–272, 1953.

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 UCSD, 2019.

Rene Caldentey, Ying Liu, and Ilan Lobel. Intertemporal pricing under minimax regret. *Operations Research*, 65(1): 104–129, 2016.

Vinicius Carrasco, Vitor Farinha Luz, Nenad Kos, Matthias Messner, Paulo Monteiro, and Humberto Moreira. Optimal selling mechanisms under moment conditions. *Journal of Economic Theory*, 177: 245–279, 2018.

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.

Yeon-Koo Che and Ian Gale. The optimal mechanism for selling to a budget-constrained buyer. *Journal of Economic Theory*, 92: 198–233.

Yiwei Chen and Vivek Farias. Robust dynamic pricing with strategic customers. *Mathematics of Operations Research*, 43(4): 1119–1142, 2018.

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

Daniele Condorelli and Balázs Szentes. Information design in the hold-up problem. *Journal of Political Economy*, Forthcoming, 2018.

John Conlisk. A peculiar example of temporal price discrimination. *Economics Letters*, 15: 121–126, 1984.

Pascal Courty and Hao Li. Sequential screening. *Review of Economic Studies*, 67: 697–717, 2000.

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

Songzi Du. Robust mechanisms under common valuation. *Econometrica*, 86(5): 1569–1588, 2018.

Piotr Dworczak and Giorgio Martini. The simple economics of optimal persuasion. *Journal of Political Economy*, 127(5): 1993–2048, 2019.

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

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.

Itay Goldstein and Yaron Leitner. Stress tests and information disclosure. *Journal of Economic Theory*, 177: 34–69, 2018.

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

Nicolas Inostroza and Alessandro Pavan. Persuasion in global games with application to stress testing. Working Paper, Northwestern University and CEPR, 2019.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101: 2590–2615, 2011.

Shaowei Ke and Qi Zhang. Randomization and Ambiguity Aversion. Working Paper, University of Michigan, 2019.

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

Michael Landsberger and Isaac Meilijson. Intertemporal price discrimination and sales strategy under incomplete information. *Rand Journal of Economics*, 16(3):424-430, 1985.

Hao Li and Xianwen Shi. Discriminatory information disclosure. *American Economic Review*, 107(11): 3363–3385, 2017.

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

Chenghu Ma and Wing-Keung Wong. Stochastic dominance and risk measure: A decision-theoretic foundation for VaR and C-VaR. *European Journal of Operational Research*, 207: 927-935, 2010.

Miltos Makris and Ludovic Renou. Information design in multi-stage games. Working Paper, Queen Mary University and University of Kent, 2019.

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.

Doron Ravid, Anne-Katrin Roesler, and Balázs Szentes. Learning before trading: on the inefficiency of ignoring free information. Working Paper, University of Chicago, University of Toronto and London School of Economics, 2019.

Frank Riedel. Optimal stopping with multiple priors. *Econometrica*, 77(3): 857-908, 2009.

John Riley and Richard Zeckhauser. Optimal selling strategies: When to haggle, when to hold firm. *Quarterly Journal of Economics*, 98(2): 267–289, 1983.

Anne-Katrin Roesler and Balázs Szentes. Buyer-optimal learning and monopoly pricing. *American Economic Review*, 107(5): 2072-2080, 2017.

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

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

Stefan Terstiege and Cédric Wasser. Buyer-optimal extensionproof information. Working Paper, Maastricht University and University of Bonn, 2019.

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

Online Appendix for "Informational Robustness in Intertemporal Pricing"

B. PROOFS FOR THE PRICE-INDEPENDENT MODEL

In this appendix, we first review the solution to the one-period model without price-dependence. 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 proof of Theorem 2.

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.:

$$
D(x) = \begin{cases} 0 & x < W \\ \frac{\log \frac{x}{W}}{\log \frac{S}{W}} & x \in [W, S) \\ 1 & x \geq S \end{cases}
\tag{23}
$$

Recall from the main text that W and B are parameters for the Roesler-Szentes information structure; see (5). In the above we have an additional parameter S, which is characterized by $S \in [W, B]$ and

$$
\int_0^S F_W^B(v) \, dv = \int_0^S F(v) \, dv
\tag{24}
$$

where F_W^B is the distribution of posterior expected values under the Roesler-Szentes worst-case information structure. To explain where S comes from, note that the LHS in (24) must not exceed the RHS *for all* S because F is a mean-preserving spread of F_W^B (Rothschild and Stiglitz (1970)). When W is smallest possible, such a constraint must bind at some S.[40]

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

---

[40]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, the Roesler-Szentes information structure has the property that any buyer with true value $v \leq S$ has posterior expected value at most S. Likewise any buyer with true value $v > S$ has posterior expected value greater than S. This fact is also pointed out by Ravid, Roesler and Szentes (2019), who call S a "separating price" in their setting.

buyer's posterior expected values. Then the seller's profit is

$$\Pi = \int_W^S p(1 - \tilde{F}(p)) \, dD(p) = \frac{1}{\log \frac{S}{W}} \int_W^S (1 - \tilde{F}(p)) \, dp \geq \frac{1}{\log \frac{S}{W}} \left(S - W - \int_0^S \tilde{F}(p) \, dp \right)$$

$$\geq \frac{1}{\log \frac{S}{W}} \left(S - W - \int_0^S F(p) \, dp \right) = \frac{1}{\log \frac{S}{W}} \left(S - W - \int_0^S F_W^B(p) \, dp \right) = W.$$

The second inequality follows because F is a mean-preserving spread of $\tilde{F}$. The next equality uses (24), and the last equality uses (5).

B.2. Proof of Lemma 3 and Theorem 2

As discussed in the main text, Theorem 2 follows from Lemma 3. So we focus on proving the lemma. The proof is broken down into several steps.

In this proof, we start with a general (price-independent) dynamic information structure $\mathcal{I}$. We use it to construct an information structure that only provides information to the buyer in period 1, while delivering lower profit to the seller.

Step 1: Cutoff prices and purchase probabilities. By assumption, the buyer's expected value follows a martingale process $v_1, v_2, \ldots$ that is autonomous (independent of the realized constant price). We define a sequence of *cutoff prices* adapted to the v-process:

$$v_t - r_t = \max_{\tau > t} \mathbb{E}[\delta^{\tau - t}(v_\tau - r_t) \mid v_1, \ldots, v_t];$$

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

In case T is finite, we extend these definitions to $t > T$ by letting $r_t = v_t = v_T$ and $q_t = q_T$.

These cutoff prices are dual concepts of cutoff values defined in Appendix A. In particular, sale occurs in period t precisely when the random constant price p belongs to $[q_{t-1}, q_t)$. Moreover, whenever $q_t = r_t \geq q_{t-1}$ we have the following analogue of Lemma 6:

$$v_t = \mathbb{E}\left[\sum_{s \geq t} (1 - \delta)\delta^{s-t} q_s \mid v_1, \ldots, v_t \right]. \tag{25}$$

Step 2: Profit decomposition. Suppose the seller draws a random price p from some c.d.f. H. Let

$$\pi(q) = \int_0^q p \, dH(p)$$

2

denote the one-period profit from a buyer whose value is q. Then we can compute total profit to be

$$\Pi = \mathbb{E}\left[\sum_{t \geq 1} \delta^{t-1} \int_{q_{t-1}}^{q_t} p \, dH(p)\right]$$

$$= \mathbb{E}\left[\sum_{t \geq 1} \delta^{t-1}(\pi(q_t) - \pi(q_{t-1}))\right]$$

$$= \mathbb{E}\left[\sum_{t \geq 1}(1 - \delta)\delta^{t-1}\pi(q_t)\right].$$

Step 3: Replacement. Given the q_t process from Step 1, define $\tilde{v}$ to be the random variable that is *equal to q_t with probability* $(1 - \delta)\delta^{t-1}$; let $\tilde{F}$ be the resulting distribution of $\tilde{v}$. Step 2 implies that profit under the dynamic information process is also the profit in one period facing the value distribution $\tilde{F}$. To complete the proof, it suffices to show that $\tilde{F}$ is the distribution of posterior expected values under prior F and some static information structure; this is, F is a mean-preserving spread of $\tilde{F}$ (see Rothschild-Stiglitz (1970)).

To do this, observe that F is a mean-preserving spread of the distribution of $v_\infty = \lim_{t \to \infty} v_t$. So it suffices to show that the latter distribution is a mean-preserving spread of $\tilde{F}$, i.e., the distribution of v_∞ should be second-order stochastically dominated by the (suitably averaged) distribution of q_t. For each real number x, let γ be a stopping time adapted to the v-process such that q_γ first exceeds x. Then

$$\mathbb{E}\left[\sum_{t \geq 1}(1 - \delta)\delta^{t-1}(q_t - x)^+\right] = \mathbb{E}\left[\delta^{\gamma-1}\sum_{t \geq \gamma}(1 - \delta)\delta^{t-\gamma}(q_t - x)\right]$$

$$= \mathbb{E}\left[\delta^{\gamma-1}(v_\gamma - x)\right]$$

$$\leq \mathbb{E}[(v_\infty - x)^+],$$

where we use y^+ to denote $\max\{y, 0\}$. The first equality follows from the definition of γ and the fact that q_t increases in t. The second equality holds by (25), which can be applied here because $q_\gamma > x \geq q_{\gamma-1}$ by definition of γ; note that it also trivially holds when $\gamma = \infty$, meaning $q_T < x$. To show the last inequality, we have $v_\gamma - x \leq (v_\gamma - x)^+ \leq \mathbb{E}[(v_\infty - x)^+ \mid v_1, \ldots, v_\gamma]$ by martingale property of the v-process and convexity of the positive part function.

Since $\mathbb{E}\left[\sum_{t \geq 1}(1 - \delta)\delta^{t-1}(q_t - x)^+\right] \leq \mathbb{E}[(v_\infty - x)^+]$ for each x, and $\mathbb{E}\left[\sum_{t \geq 1}(1 - \delta)\delta^{t-1}q_t\right] = \mathbb{E}[v_1] = \mathbb{E}[v]$ by (25), we conclude SOSD as desired. Lemma 3 and Theorem 2 then follow.

C. PROOFS FOR OTHER EXTENSIONS

C.1. Proof of Proposition 4

If information only arrives once, we will show that a seller who sets prices $p_2 = p^*$ and p_1 slightly larger than p^* can guarantee strictly more than Π^*. The proof considers two cases (information either in the first period or second):

Case 1: Information in period one. Let $\tilde{F}$ denote the distribution of posterior expected values given the static information structure. Then profit can be computed as

$$
\begin{aligned}
\Pi &= p_1(1 - \tilde{F}(v_1)) + \delta p_2(\tilde{F}(v_1) - \tilde{F}(v_2)) \\
&= (1 - \delta)v_1(1 - \tilde{F}(v_1)) + \delta v_2(1 - \tilde{F}(v_2)),
\end{aligned}
\tag{26}
$$

where $v_1 = \frac{p_1 - \delta p_2}{1 - \delta}$ and $v_2 = p_2$ are the threshold values for buying in period one and period two, respectively. Since F is a mean-preserving spread of $\tilde{F}$, we have

$$
\int_0^x F(s)\, ds \geq \int_0^x \tilde{F}(s)\, ds, \quad \forall 0 \leq x \leq 1.
$$

By our choice, $v_2 = p_2 = p^*$ and v_1 is slightly larger than p^*. Then for all $x > v_1 > p^*$ the above inequality implies a joint upper-bound on $\tilde{F}(v_1)$ and $\tilde{F}(v_2)$ as follows:

$$
\int_0^x F(s)\, ds \geq \int_0^x \tilde{F}(s)\, ds \geq (v_1 - p^*)\tilde{F}(p^*) + (x - v_1)\tilde{F}(v_1),
\tag{27}
$$

where the second inequality holds by monotonicity of the c.d.f. $\tilde{F}$.

In particular, let us choose $x = L^{-1}(p^*) = F^{-1}(G(p^*))$. Note that $G(p^*) > 0$ ensures $x > p^* > \underline{v}$, so for p_1 close to p^* we indeed have $v_1 \in (p^*, x)$. Moreover, $p^* = \frac{1}{F(x)} \int_0^x s\, dF(s)$ and so

$$
\int_0^x F(s)\, ds = xF(x) - \int_0^x s\, dF(s) = xF(x) - p^*F(x) = (x - p^*)G(p^*).
$$

Combined with (27), we deduce the following inequality:

$$
\tilde{F}(v_1) - G(p^*) \leq \frac{v_1 - p^*}{x - v_1} \cdot (G(p^*) - \tilde{F}(p^*)).
\tag{28}
$$

Plugging into the objective function (26), we conclude that for v_1 sufficiently close to p^* and $\epsilon > 0$

sufficiently small compared to $v_1 - p^*$ (see below), it holds that

$$
\begin{aligned}
\Pi &= (1 - \delta)v_1(1 - \tilde{F}(v_1)) + \delta p^*(1 - \tilde{F}(p^*)) \\
&\geq (1 - \delta)p^*(1 - \tilde{F}(v_1)) + \delta p^*(1 - \tilde{F}(p^*)) + \epsilon \\
&= p^* \left[1 - G(p^*) + \delta(G(p^*) - \tilde{F}(p^*)) - (1 - \delta)(\tilde{F}(v_1) - G(p^*)) \right] + \epsilon \\
&\geq p^*(1 - G(p^*)) + \epsilon \\
&= \Pi^* + \epsilon.
\end{aligned}
$$

The inequality in the second line holds whenever $\epsilon \leq (v_1 - p^*)(1 - \tilde{F}(v_1))$. As $v_1 \to p^*$, we have $\limsup \tilde{F}(v_1) \leq G(p^*) < 1$ from (28). Thus we are able to choose some $\epsilon > 0$ (depending on v_1) that satisfies this inequality. As for the inequality in the penultimate line above, it holds because $\frac{\delta}{1-\delta} \geq \frac{v_1 - p^*}{x - v_1}$ and $G(p^*) - \tilde{F}(p^*) \geq 0$, the latter of which follows from (28) and $\tilde{F}(v_1) \geq \tilde{F}(p^*)$.

Hence when information only arrives in the first period, the seller guarantees more than Π^*.

Case 2: Information in period two. Suppose instead that the buyer only receives a signal in the second period. If the information structure is such that the buyer prefers to purchase in period one, profit clearly increases to p_1. Below we focus on the situation where information in the second period makes the buyer willing to delay. Then incentive compatibility requires that

$$
\mathbb{E}[v] - p_1 \leq \delta \times \textit{expected buyer surplus in period two}
$$

Since $\delta < 1$ and p_1 is slightly larger than p_2, buyer surplus in period two is greater than (and bounded away from) $\mathbb{E}[v] - p_2$, which is the surplus under the worst-case threshold information structure against price p_2. Since this worst-case scenario maximizes buyer surplus subject to probability of sale being equal to $1 - G(p_2)$, we deduce that actual probability of sale in period two must be greater than (and bounded away from) $1 - G(p_2)$.

To proceed with the analysis, we assume without loss that there is exactly one signal $\bar{s}$ in the second period that recommends the buyer to purchase. Then we can rewrite the incentive compatibility condition as

$$
\mathbb{E}[v] - p_1 \leq \delta \cdot \mathbb{P}[\bar{s}] \cdot (\mathbb{E}[v \mid \bar{s}] - p_2). \tag{29}
$$

Since the probability of sale exceeds $1 - G(p_2)$, the expected value upon seeing $\bar{s}$ is less than (and bounded away from) the average value conditional on value above the lowest $G(p_2)$-percentile.

5

This average value is exactly $\frac{\mathbb{E}[v]-p_2 G(p_2)}{1-G(p_2)}$. Thus for some $\eta > 0$ independent of p_1, we have

$$\mathbb{E}[v \mid \bar{s}] - p_2 \leq \frac{\mathbb{E}[v] - p_2 G(p_2)}{1 - G(p_2)} - \eta - p_2 = \frac{\mathbb{E}[v] - p_2}{1 - G(p_2)} - \eta.$$

Therefore we have the following profit lower-bound:

$$\Pi = \delta \cdot \mathbb{P}[\bar{s}] \cdot p_2 \geq (\mathbb{E}[v] - p_1) \cdot \frac{p_2}{\mathbb{E}[v \mid \bar{s}] - p_2} \geq \frac{(\mathbb{E}[v] - p_1)p_2}{\frac{\mathbb{E}[v] - p_2}{1 - G(p_2)} - \eta},$$

where the first inequality uses the IC constraint (29).

As $p_1 \to p_2 = p^*$, the RHS above is larger than $p_2(1 - G(p_2)) = \Pi^*$, completing the proof of the proposition.

C.2. Proof of Proposition 5

We first assume the truth of the Replacement Lemma. Let $\tilde{F}$ denote the distribution of posterior valuations arising from an arbitrary static information structure. Then the seller's total profit under this information structure can be written as:

$$(1 - \delta) \cdot \Pi^C(\delta, T) = \min_{\tilde{F}} \sum_{t=1}^{T} (1 - \delta)\delta^{t-1} p_t \cdot (1 - \tilde{F}(p_t)), \tag{30}$$

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}.

To achieve this approximation, we equate the c.d.f. 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 $p_1, \ldots, p_T$ are densely distributed on the interval (W, S). Hence their distribution converges to $D(\cdot)$, which proves the proposition. We turn to Lemma 4.

Proof of Lemma 4. The proof strategy is analogous to Lemma 1, with difference due to the interdependence of information and values across buyers. More precisely, 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. Under this replacement, each buyer a (i.e., the buyer who

arrives in period a) either purchases in period a at the price p_a, or never. Since prices increase over time, we deduce that each buyer purchases at lower prices. If we can further ensure that the discounted probability of sale to each buyer is lower than the original information structure, then profit is necessarily decreased. The construction is broken down into several steps below.

Step 1: Stopping times and critical buyers. We first define a family of random variables $\{\tau(a)\}_{a=1}^{T}$ adapted to the process of signals under the original information structure $\mathcal{I}$. Each $\tau(a)$ denotes the optimal stopping time of buyer a, i.e., this buyer finds it optimal to purchase in period $\tau(a)$ given signal realizations $s_1, \ldots, s_{\tau(a)}$. Note that $\tau(a) \geq a$, since buyer a can only purchase starting from that period. Due to public signals, we additionally have $\tau(a+1) = \tau(a)$ whenever $\tau(a) > a$; this equality captures the observation that if buyer a delays purchase, then in every future period she faces the same problem as the next buyer $a + 1$.

Given these stopping times $\{\tau(a)\}$, we define a "critical set" $C = \{j_1, j_2, \cdots, j_n, T+1\}$ of buyers as follows. To begin, j_1 is the first buyer who delays purchase (formally, $j_1 = \min\{a : \tau(a) > a\}$). Next, j_2 is the first buyer after $\tau(j_1)$ that delays purchase (i.e., $j_2 = \min\{a > \tau(j_1) : \tau(a) > a\}$). So on and so forth, until we have reached some j_n such that every buyer arriving after period $\tau(j_n)$ purchases immediately upon arrival (along the history of signals being considered). To simplify some of the later exposition, we include a hypothetical buyer $j = T + 1$ into the critical set C, and define $\tau(T + 1) = \infty$. We note that the critical buyers and their stopping times pin down the stopping behavior of all the buyers: Specifically, buyers $a \in [j_m, \tau(j_m))$ all delay purchase to period $\tau(j_m)$, whereas each buyer $a \in [\tau(j_m), j_{m+1})$ purchases immediately in period a (this also applies to $a < j_1$ and $a \geq \tau(j_n)$).[41]

Step 2: Replacement. Now we are ready to construct the replacement information structure $\mathcal{I}'$. We take the signal set to be $S^* = \{0, 1, \ldots, T\}$, where the signal realization $s^* = i$ represents nature's recommendation that the first i buyers purchase upon arrival and that other buyers do not purchase (buyer obedience will be verified later). To generate these signals, we fix the true value v and draw any history of signals $s_1, \ldots, s_T$ under the original information structure. Then, in the replacement information structure, the probabilities of different signals are given by

$$\mathbb{P}[s^* = i \mid s_1, \ldots, s_T] = \begin{cases} \delta^{\sum_{k<m} \tau(j_k) - j_k} \cdot \left(1 - \delta^{\tau(j_m) - j_m}\right), & \text{if } i = j_m - 1 \text{ for some } j_m \in C(s_1, \ldots, s_T); \\ 0, & \text{otherwise.} \end{cases}$$

Step 3: Lower probability of sale. Assuming that buyers follow nature's recommendation *not* to purchase, we now show that this signal structure leads to lower discounted probability of sale

---

[41] 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.

7

to each buyer. From the above specification of probabilities, we see that for each buyer a, the probability of receiving $s^* \geq a$ is

$$\mathbb{P}[s^* \geq a \mid s_1, \ldots, s_T] = \delta^{\sum_{k \leq m} \tau(j_k) - j_k},$$

where j_m is the last critical buyer up to and including a. Now notice that $\tau(j_m) - j_m \geq \tau(a) - a$, because if $a \in [j_m, \tau(j_m))$ then $\tau(a) = \tau(j_m)$ so the RHS is smaller, and if $a \in [\tau(j_m), j_{m+1})$ then $\tau(a) = a$ and the RHS is again smaller. Hence, the probability of receiving $s^* \geq a$ is

$$\delta^{\sum_{k \leq m} \tau(j_k) - j_k} \leq \delta^{\tau(j_m) - j_m} \leq \delta^{\tau(a) - a}.$$

It follows that buyer a's discounted purchase probability in the replacement information structure is at most $\delta^a \cdot \delta^{\tau(a) - a} \leq \delta^{\tau(a)}$, which is the probability under the original information structure.

Step 4: Buyer obedience. Finally, we verify that if buyer a receives signal $s^* = i < a$, then she optimally follows nature's recommendation not to purchase the object. That is, we need to show that her expected value given $s^* = i$ is at most p_a. Since all buyers have the same expectation (under $\mathcal{I}'$) and prices are increasing over time, it is sufficient to consider $a = i + 1$. We will prove a stronger result, that conditional on the signal $s^* = a - 1$ and on any realizations $s_1, \ldots, s_a$ for which s^* is possible, expected value is at most p_a. Note that once $s_1, \ldots, s_a$ are fixed, then so are the critical buyers up to and including a (because whether a buyer delays purchase only depends on past information). Without loss we assume a is critical, since otherwise $s^* = a - 1$ does not occur.

Let $a = j_m$ be a critical buyer for the fixed signal history $s_1, \ldots, s_a$. Then for each $k < m$, the identity of the critical buyer j_k and her stopping time $\tau(j_k) < a$ are fixed. Below we show $\mathbb{E}[v \mid s_1, \ldots, s_a, s^* = a - 1] \leq p_a$. For this we follow the proof of Lemma 1 in Appendix A. For $t > a$, define $y_t = \mathbb{E}[v \mid s_1, \ldots, s_a, \tau(a) = t]$ to be buyer a's expected value conditional on the history $s_1, \ldots, s_a$ and conditional on being recommended to purchased in period t, under the original information structure. Since purchasing in period $\tau(a)$ is by definition better than purchasing in period a, we have

$$\mathbb{E}\left[v \mid s_1, \ldots, s_a\right] - p_a \leq \mathbb{E}\left[\delta^{\tau(a) - a} \cdot \left(y_{\tau(a)} - p_{\tau(a)}\right) \mid s_1, \ldots, s_a\right].$$

Note that $p_{\tau(a)} \geq p_a$, we thus have

$$\mathbb{E}\left[v \mid s_1, \ldots, s_a\right] - p_a \leq \mathbb{E}\left[\delta^{\tau(a) - a} \cdot \left(y_{\tau(a)} - p_a\right) \mid s_1, \ldots, s_a\right].$$

Rearranging and using $\mathbb{E}[v \mid s_1, \ldots, s_a] = \mathbb{E}[y_{\tau(a)} \mid s_1, \ldots, s_a]$ (Law of Iterated Expectations), we

can obtain

$$p_a \geq \frac{\mathbb{E}\left[(1 - \delta^{\tau(a)-a}) \cdot y_{\tau(a)} \mid s_1, \ldots, s_a\right]}{\mathbb{E}\left[1 - \delta^{\tau(a)-a} \mid s_1, \ldots, s_a\right]}. \tag{31}$$

This is the analogue of (14) in the current setting.

Now recall that $a = j_m$ is a critical buyer, and the previous critical buyers $j_1, \ldots, j_{m-1}$ as well as their stopping times $\tau(j_1), \ldots, \tau(j_{m-1})$ are determined by the signal history $s_1, \ldots, s_a$. Thus, $\delta^{\sum_{k<m} \tau(j_k)-j_k}$ is a constant once we condition on $s_1, \ldots, s_a$. (31) then implies that

$$p_a \geq \frac{\mathbb{E}\left[\delta^{\sum_{k<m} \tau(j_k)-j_k}(1 - \delta^{\tau(a)-a}) \cdot y_{\tau(a)} \mid s_1, \ldots, s_a\right]}{\mathbb{E}\left[\delta^{\sum_{k<m} \tau(j_k)-j_k}(1 - \delta^{\tau(a)-a}) \mid s_1, \ldots, s_a\right]}, \tag{32}$$

By construction, the term $\delta^{\sum_{k<m} \tau(j_k)-j_k}(1-\delta^{\tau(a)-a})$ is precisely the probability that $s^* = j_m - 1 = a - 1$, conditional on the entire signal history $s_1, \ldots, s_T$. We can also write it as the conditional probability $\mathbb{P}[s^* = a - 1 \mid s_1, \ldots, s_a, \tau(a)]$, since the remaining signals $s_{a+1}, \ldots, s_T$ affect the distribution of s^* only via the stopping time $\tau(a)$.

Therefore, (32) can be rewritten as

$$p_a \geq \frac{\mathbb{E}\left[\mathbb{P}[s^* = a - 1 \mid s_1, \ldots, s_a, \tau(a)] \cdot y_{\tau(a)} \mid s_1, \ldots, s_a\right]}{\mathbb{E}\left[\mathbb{P}[s^* = a - 1 \mid s_1, \ldots, s_a, \tau(a)] \mid s_1, \ldots, s_a\right]}. \tag{33}$$

By the Law of Iterated Expectations again, the denominator is simply the conditional probability $\mathbb{P}[s^* = a-1 \mid s_1, \ldots, s_a]$. On the other hand, $y_{\tau(a)}$ in the numerator is the conditional expectation $\mathbb{E}[v \mid s_1, \ldots, s_a, \tau(a)]$, which is also $\mathbb{E}[v \mid s_1, \ldots, s_a, \tau(a), s^* = a - 1]$ since s^* does not provide additional information about v beyond $s_1, \ldots, s_a$ and $\tau(a)$.

So we can further rewrite (33) as

$$p_a \geq \frac{\mathbb{E}\left[\mathbb{P}[s^* = a - 1 \mid s_1, \ldots, s_a, \tau(a)] \cdot \mathbb{E}[v \mid s_1, \ldots, s_a, \tau(a), s^* = a - 1] \mid s_1, \ldots, s_a\right]}{\mathbb{P}[s^* = a - 1 \mid s_1, \ldots, s_a]}.$$

As the denominator $\mathbb{P}[s^* = a - 1 \mid s_1, \ldots, s_a]$ is a constant conditional on $s_1, \ldots, s_a$, we can absorb it into the numerator:

$$p_a \geq \mathbb{E}\left[\frac{\mathbb{P}[s^* = a - 1 \mid s_1, \ldots, s_a, \tau(a)]}{\mathbb{P}[s^* = a - 1 \mid s_1, \ldots, s_a]} \cdot \mathbb{E}[v \mid s_1, \ldots, s_a, \tau(a), s^* = a - 1] \mid s_1, \ldots, s_a\right]$$

Using the identity $\frac{\mathbb{P}[s^*=a-1 \mid s_1,\ldots,s_a,\tau(a)]}{\mathbb{P}[s^*=a-1 \mid s_1,\ldots,s_a]} = \frac{\mathbb{P}[\tau(a) \mid s_1,\ldots,s_a,s^*=a-1]}{\mathbb{P}[\tau(a) \mid s_1,\ldots,s_a]}$, we arrive at

$$p_a \geq \mathbb{E}\left[\frac{\mathbb{P}[\tau(a) \mid s_1, \ldots, s_a, s^* = a - 1]}{\mathbb{P}[\tau(a) \mid s_1, \ldots, s_a]} \cdot \mathbb{E}[v \mid s_1, \ldots, s_a, \tau(a), s^* = a - 1] \mid s_1, \ldots, s_a\right]$$

9

We can now show that the RHS above is equal to $\mathbb{E}[v \mid s_1, \ldots, s_a, s^* = a - 1]$, which will complete the proof. Indeed, the conditional expectation on the RHS is taken with respect to the distribution of $\tau(a)$ conditional on $s_1, \ldots, s_a$. It can thus be rewritten as

$$\sum_{t > a} \mathbb{P}[\tau(a) = t \mid s_1, \ldots, s_a] \cdot \frac{\mathbb{P}[\tau(a) = t \mid s_1, \ldots, s_a, s^* = a - 1]}{\mathbb{P}[\tau(a) = t \mid s_1, \ldots, s_a]} \cdot \mathbb{E}[v \mid s_1, \ldots, s_a, \tau(a) = t, s^* = a-1],$$

which is just

$$\sum_{t > a} \mathbb{P}[\tau(a) = t \mid s_1, \ldots, s_a, s^* = a - 1] \cdot \mathbb{E}[v \mid s_1, \ldots, s_a, \tau(a) = t, s^* = a - 1].$$

This is precisely $\mathbb{E}[v \mid s_1, \ldots, s_a, s^* = a - 1]$. $\qquad\square$

D. DISTRIBUTIONAL UNCERTAINTY

Our main model assumes that the seller knows the prior value distribution F. This assumption enables us to focus on "informational uncertainty", which influences how the buyer's expected value evolves over time and is thus relevant for the seller's *dynamic* pricing strategy. However, our results also have implications for sellers who additionally face "distributional uncertainty", i.e., uncertainty about the distribution F. We discuss this extension below.

Formally, we consider here a seller who thinks the prior distribution F is chosen (by nature) from a family of distributions $\mathcal{F}$. The buyer knows F to begin with, but can potentially receive more information about her value via some process $\mathcal{I}$, just as in our main model. The seller commits to a pricing strategy to maximize his worst-case profit, where the worst case is evaluated with respect to all possible priors $F \in \mathcal{F}$ and all information processes $\mathcal{I}$.

A special case of such a model is when $\mathcal{F}$ is the set of all distributions supported on an interval $[a, b]$, with a given mean $\mu \in (a, b)$. Let F_0 be the distribution supported on the two extreme values a, b, with mean μ. Then any distribution $F \in \mathcal{F}$ is a *mean-preserving contraction* of F_0, which means that F can be thought of as the distribution of posterior expected values under the prior F_0 and some static information structure. As a result, nature could replicate the choice of any prior F by choosing F_0 as the prior, and providing this information structure in period 1. This suggests that F_0 is the *worst-case prior distribution* for the seller (intuitively because it leaves the most amount of residual uncertainty). Given that the prior is now "fixed," we can apply the results for our main model to argue that a constant price path is the seller's robustly optimal strategy.

In general, we have the following result:

Proposition 7. *Suppose there exists a possible prior distribution $F_0 \in \mathcal{F}$ and a price p_0 with the following properties:*

1. *$p_0 \in \arg\max_p p(1 - G_0(p))$, where G_0 is the pressed distribution of F_0;*

2. *$G_0(p_0) \geq G(p_0)$ for the pressed distribution G of any other $F \in \mathcal{F}$.*

Then the seller's robustly optimal strategy is a constant price path of p_0, with profit guarantee $\Pi_0 = p_0(1 - G_0(p_0))$.

Proof. Clearly, the seller cannot guarantee more than Π_0 because nature can always choose F_0 as the prior. It thus remains to show that always charging p_0 guarantees profit Π_0 even when nature can choose any distribution $F \in \mathcal{F}$. This follows from the Replacement Lemma (Lemma 1), which implies that for any prior $F \in \mathcal{F}$, a constant price path of p_0 guarantees profit $p_0(1 - G(p_0)) \geq p_0(1 - G_0(p_0)) = \Pi_0$. $\square$

We can view nature's choice of the prior distribution and the seller's price as their respective strategies in a zero-sum game. From this perspective, the two conditions in Proposition 7 together imply that F_0 and p_0 constitute a *saddle point* of this game. In what follows, we demonstrate sufficient conditions on the set $\mathcal{F}$ to guarantee the existence of a saddle point.

D.1. Sufficient Condition for Proposition 7: SOSD

First, we generalize the example before Proposition 7 to show that if $F_0 \in \mathcal{F}$ is *second-order stochastically dominated* by every $F \in \mathcal{F}$, then F_0 is the worst-case prior distribution. Intuitively, when $F_0 \preceq_{SOSD} F$, it can be obtained from F by moving toward lower values and/or mean-preserving spreads. In the former case, F_0 is concentrated on lower values compared to F, and is thus a worse value distribution for profit. In the latter case, F_0 is a mean-preserving spread of F, which is also worse for profit in the presence of informational uncertainty as we discussed before.

Theorem 3. *Suppose there exists a possible prior distribution F_0 such that $F_0 \preceq_{SOSD} F$ holds for every $F \in \mathcal{F}$. Let G_0 be the pressed distribution of F_0. Then the seller's robustly optimal strategy is a constant price path of $p_0 \in \arg\max_p p(1 - G_0(p))$, with profit guarantee $\Pi_0 = p_0(1 - G_0(p_0))$.*

The formal proof follows from Proposition 7 and the lemma below, which relates *second-order* stochastic dominance between two prior distributions to *first-order* stochastic dominance between their pressed distributions. A version of this lemma appears as Theorem 2 in Ma and Wong (2010).

Lemma 7. *$F_0 \preceq_{SOSD} F$ if and only if their pressed distributions satisfy $G_0 \preceq_{FOSD} G$, i.e., $G(p) \leq G_0(p), \forall p$.*

Proof. We present the proof assuming that F and F_0 are bounded continuous distributions; the general case follows from an approximation argument. We will show that F dominates F_0 in SOSD if and only if for each $t \in (0, 1]$, the expected value of the lowest t-percentile under F is weakly higher than under F_0, which is in turn equivalent to the statement that $G(p) \leq G_0(p)$ for every p.

In one direction, suppose $F_0 \preceq_{SOSD} F$, then for each $y \in \mathbb{R}$,

$$\int_{-\infty}^{y} F(x)\, dx \leq \int_{-\infty}^{y} F_0(x)\, dx. \tag{34}$$

What we need to show is that for each $t \in [0, 1]$,

$$\int_{0}^{t} F^{-1}(q)\, dq \geq \int_{0}^{t} F_0^{-1}(q)\, dq. \tag{35}$$

This inequality clearly holds at $t = 0$, where both sides are 0. It also holds at $t = 1$, where the two sides are the unconditional expectations of F and F_0 respectively ($\mathbb{E}[F]$ is higher because F is better in SOSD). Thus we just need to check interior extreme points t. By differentiating (35) with respect to t, we only need to consider those t where $F^{-1}(t) = F_0^{-1}(t)$, which we denote by y. Then, the LHS of (35) becomes

$$\int_{0}^{F(y)} F^{-1}(q)\, dq = \int_{-\infty}^{y} x\, dF(x) = yF(y) - \int_{-\infty}^{y} F(x)\, dx = yt - \int_{-\infty}^{y} F(x)\, dx.$$

Similarly the RHS of (35) is

$$yt - \int_{-\infty}^{y} F_0(x)\, dx.$$

Hence (35) follows directly from (34).

Conversely, suppose (35) holds and we want to deduce (34). Note that (34) holds for $y \to -\infty$, where both sides are 0. It also holds for $y \to \infty$, where the difference between the RHS and LHS of (35) is $\mathbb{E}[F] - \mathbb{E}[F_0]$, which is positive because (35) holds at $t = 1$. Thus again we only need to check interior extreme points, which satisfy $F(y) = F_0(y)$. Denoting this common percentile by t, we can then reverse the above calculation and deduce (34) from (35). $\qquad \square$

D.2. Another Sufficient Condition for Proposition 7: Regularity

In cases where there does not exist a distribution F_0 that is worst in terms of SOSD, the following result provides a different sufficient condition for Proposition 7 to apply. We make use of Sion's minimax theorem to deduce that under certain assumptions, nature has a minmax prior distribution,

whereas the seller has a maxmin price.

Theorem 4. *Suppose that the set of possible priors $\mathcal{F}$ has the following properties:*

- *There exists $\overline{v} < \infty$ such that each $F \in \mathcal{F}$ is supported on $[0, \overline{v}]$ and admits a density f;*

- *$\mathcal{F}$ is closed with respect to the weak-* topology, and convex with respect to mixture;*

- *For each $F \in \mathcal{F}$ and its pressed distribution G, the function $p(1 - G(p))$ is quasi-concave (i.e., single-peaked) in p for $p \geq 0$.*

Then there exists $F_0 \in F$ and $p_0 \geq 0$ that satisfy the two conditions in Proposition 7. The seller's robustly optimal strategy is a constant price path of p_0, with profit guarantee $\Pi_0 = p_0(1 - G_0(p_0))$.

The quasi-concavity of $p(1 - G(p))$ is characterized in the following lemma:

Lemma 8. *Given F with density f and its pressed distribution G. The function $p(1 - G(p))$ is quasi-concave if and only if the function $\int_0^x F(t)dt - xF^2(x)$ crosses zero exactly once or at a single interval of points, from the above.*

A sufficient condition is that $2t - \frac{1 - F(t)}{f(t)}$ increases in t (weaker than the usual regularity condition).

Below we prove the lemma and the theorem in turn.

Proof of Lemma 8. Let $x = L^{-1}(p)$ be the value type below which the expected value is p. Since we are concerned with quasi-concavity, we can equivalently write profit as a function of x. Specifically,

$$\Pi(x) = p \cdot (1 - G(p)) = \frac{\int_0^x tf(t)\,dt}{F(x)} \cdot (1 - F(x)) = \frac{\int_0^x tf(t)\,dt}{F(x)} - \int_0^x tf(t)\,dt.$$

Taking the derivative, we obtain

$$\Pi'(x) = \frac{xf(x)F(x) - f(x)\int_0^x tf(t)\,dt}{F^2(x)} - xf(x)$$

$$= \frac{f(x)}{F^2(x)} \cdot \left(xF(x) - \int_0^x tf(t)\,dt - xF^2(x) \right)$$

$$= \frac{f(x)}{F^2(x)} \cdot \left(\int_0^x F(t)\,dt - xF^2(x) \right).$$

Thus, Π is quasi-concave/single-peaked in x if and only if $\Pi'(x)$ is first positive then negative, which is in turn equivalent to the statement that $\int_0^x F(t)dt - xF^2(x)$ crosses zero exactly once, from the above.

As for the sufficient condition, we can derive it by writing

$$\int_0^x F(t)\,dt - xF^2(x) = \int_0^x F(t)\cdot(1 - F(t) - 2tf(t))\ dt.$$

If $\frac{1-F(t)}{f(t)} - 2t$ is decreasing in t, then it is first positive then negative. This implies the entire integrand

$$h(t) := F(t)(1 - F(t) - 2tf(t)) = F(t)\cdot f(t)\cdot\left(\frac{1 - F(t)}{f(t)} - 2t\right)$$

is first positive then negative. Hence so is the function $\int_0^x h(t)\ dt$, as we desire to show. □

Proof of Theorem 4. Each distribution $F \in \mathcal{F}$ can be viewed as a continuous c.d.f. on $[0, \bar{v}]$. So the set $\mathcal{F}$ is a subset of the topological vector space of continuous functions on $[0, \bar{v}]$, equipped with the sup norm.[42] By assumption $\mathcal{F}$ is convex. $\mathcal{F}$ is also compact because for any sequence $\{F_n\} \subset \mathcal{F}$, there is a subsequence that converges weakly to some distribution F (since the space of probability distributions on an interval is weak-* compact). Since $\mathcal{F}$ is weak-* closed, the limit F belongs to $\mathcal{F}$. As F_n converges weakly to the continuous distribution F, the c.d.f. of F_n converges in the sup norm to F.[43] Hence $\mathcal{F}$ is a compact convex subset of a topological vector space.

Now, for each $F \in \mathcal{F}$ and $p \in [0, \bar{v}]$, define the (minimum) profit function

$$\Pi(F, p) = p(1 - G_F(p)),$$

where G_F is the pressed distribution of F. For fixed F, this function is clearly continuous in p and also quasi-concave by assumption. For fixed p, this function is continuous and quasi-convex in F, as we show below. Thus, we can apply Sion's minimax theorem to deduce that

$$\min_{F\in\mathcal{F}}\ \max_{p\in[0,\bar{v}]}\ \Pi(F, p) = \max_{p\in[0,\bar{v}]}\ \min_{F\in\mathcal{F}}\ \Pi(F, p).$$

Let this maxmin/minmax profit be Π_0, then there exists $F_0 \in \mathcal{F}$ such that $\Pi(F_0, p) \leq \Pi_0$ for all p, and there exists p_0 such that $\Pi(F, p_0) \geq \Pi_0$ for all $F \in \mathcal{F}$. Hence $\Pi(F_0, p_0) = \Pi_0$ and the pair (F_0, p_0) constitutes a saddle point. That precisely implies $p_0 \in \arg\max_p p(1 - G_0(p))$, and $G_0(p_0) \geq G(p_0)$ for any other pressed distribution G of any $F \in \mathcal{F}$, as we desire to show.

It remains to verify the continuity and quasi-convexity of $\Pi(F, p)$ as a function of F. For

---

[42]We restrict to continuous distributions so that the space of c.d.f. can be embedded in a topological vector space.
[43]The argument is as follows: For each $\epsilon > 0$, uniform continuity of F allows us to choose $\delta \leq \epsilon$ such that $|F(x) - F(y)| \leq \epsilon$ whenever $|x - y| \leq \delta$. Next, recall that weak convergence is equivalent to convergence in the Lévy metric. Thus for this δ, there exists N such that $F(x - \delta) - \delta \leq F_n(x) \leq F(x + \delta) + \delta$ for all x and $n \geq N$. It follows that $F_n(x) \leq F(x + \delta) + \delta \leq F(x) + \epsilon + \delta \leq F(x) + 2\epsilon$, and similarly $F_n(x) \geq F(x) - 2\epsilon$ for all x and $n \geq N$. Hence $F_n \to F$ uniformly.

quasi-convexity, we need to show that for any $q, \lambda \in [0, 1]$, if $G_{F_1}(p), G_{F_2}(p) \geq q$ (so that $\Pi(F_1, p), \Pi(F_2, p) \leq p(1 - q)$), then $G_F(p) \geq q$ also holds for the mixture distribution $F = \lambda F_1 + (1 - \lambda)F_2$. By definition, the lowest $G_{F_1}(p)$-percentile of the distribution F_1 has expected value p. Thus the condition $G_{F_1}(p) \geq q$ tells us that the lowest q-percentile of F_1 has expected value at most p, and the same holds for the distribution F_2. By mixing the lowest q-percentile from F_1 with that from F_2, we know that in the distribution F, *some* fraction q of types has expected value at most p. Hence, the lowest q-percentile of F has expected value at most p, which then implies that $G_F(p) \geq q$.

As for continuity, we need to show that if $F_n \to F$ in the sup norm, then $G_{F_n}(p) \to G_F(p)$ for each p. We first show $G_F(p) \geq \limsup_{n \to \infty} G_{F_n}(p)$. By passing to a subsequence, it suffices to show that for any $\alpha > 0$, if each $G_{F_n}(p) \geq \alpha$ then $G_F(p) \geq \alpha$ also holds. Specifically, the condition $G_{F_n}(p) \geq \alpha$ implies the lowest α-percentile of each F_n has expected value at most p. That is,

$$\int_0^{F_n^{-1}(\alpha)} x \, dF_n(x) \leq \alpha p.$$

Applying integration by parts to the LHS, we obtain

$$\int_0^{F_n^{-1}(\alpha)} (\alpha - F_n(x)) \, dx \leq \alpha p.$$

Fixing any $\epsilon > 0$, then uniform convergence of F_n to F gives $F_n(x) \leq F(x) + \epsilon$ for all large n and all x. Moreover, for large n we have $F_n^{-1}(\alpha) \geq F^{-1}(\alpha - \epsilon)$. Thus, the preceding inequality implies

$$\int_0^{F^{-1}(\alpha - \epsilon)} (\alpha - \epsilon - F(x)) \, dx \leq \alpha p.$$

Applying integration by parts again, we deduce for large n

$$\int_0^{F^{-1}(\alpha - \epsilon)} x \, dF(x) \leq \alpha p.$$

Letting $\epsilon \to 0$ we thus conclude that the lowest α-percentile of F has expected value at most p, so that $G_F(p) \geq \alpha$ as desired.

A completely symmetric argument shows $G_F(p) \leq \liminf_{n \to \infty} G_{F_n}(p)$. Hence the profit function $\Pi(F, p)$ is continuous in F, finishing the proof. □

E. BUYER UNCERTAINTY

One may wonder why the buyer in our model is so much better informed than the seller. In particular, this issue may appear more salient in our dynamic setting (than models of robust static mechanism design), since we have assumed that the buyer knows all future information structures.

Allowing for the buyer to face non-Bayesian uncertainty over information arrival leads to technical difficulties related to how these beliefs update over time. Developing a general theory of dynamic non-Bayesian updating is beyond the scope of this paper. However, we make two simple observations regarding how our results would change if the buyer herself could face uncertainty over the information process, evaluating her surplus assuming the worst-case information process. For illustration, we focus on a finite horizon T and deterministic price paths throughout.

Our first observation is that, without any restriction on how much uncertainty the buyer faces, nature can hold the seller down to *zero* profit. Intuitively, if the buyer only observes signal realizations but does not understand the information structure generating these realizations, then she could always expect her value to be low in the worst case. For a specific example, suppose that nature can choose one of two possible information structures in each period t. One of these information structures generates realization $s_t = 1$ if $v \geq p_t$ and $s_t = 0$ otherwise, while the other information structure generates $s_t = 0$ if $v \geq p_t$ and $s_t = 1$ otherwise. Faced with such uncertainty, the buyer receiving any realized s_t expects her value to be below p_t in the worst case, and thus does not purchase.

To rule out such a situation, we next consider a natural restriction on buyer uncertainty, imposing that at each period t, the buyer knows the information structure *in that period* even though she faces uncertainty over *future* information. Specifically, given the seller's prices, the interaction consists of the following:

- Nature chooses an information process $\mathcal{I} = (I_t)_{t=1}^T$ with $I_t : V \times S^{t-1} \to \Delta(S_t)$ for each t. This is *not* known to the buyer.

- The buyer's value is drawn, with $v \sim F$.

- In each period t, the buyer learns the information structure I_t in that period that maps $V \times S^{t-1}$ to $\Delta(S_t)$, and also observes a signal realization s_t.

- Based on the history of information structures and signal realizations, the buyer forms a Bayesian posterior about her value.

- Given her belief, the buyer decides whether or not to purchase in period t at the price p_t, assuming that if she does not purchase, future information structures will be worst possible for her expected payoff.

We assume the buyer is sophisticated; that is, she knows that if she does not purchase in period t, her period $t+1$ self will again assume the worst information structure for the future and expect the period $t+2$ self to behave in the same way, so on and so forth.

As we show below, the buyer's optimal behavior under any information process can be determined by backward induction. Given this, the seller chooses prices to maximize his worst-case profit across different information processes.

Lemma 9. *Suppose the buyer faces uncertainty about future information structures and is sophisticated. Given any sequence of prices $p_1, \ldots, p_T$ and any (partial) history of signals $s_1, \ldots, s_t$, the buyer optimally purchases in period t if and only if*

$$\mathbb{E}[v \mid s_1, \ldots, s_t] - p_t > \delta^{\tau-t} \left(\mathbb{E}[v \mid s_1, \ldots, s_t] - p_\tau \right), \quad \forall \tau \in \{t+1, \ldots, T, \infty\}, \qquad (36)$$

where the RHS is interpreted to be 0 for $\tau = \infty$.

Proof. We prove this by backward induction. The case where $t = T$ is straightforward, as in this case the buyer purchases if and only if her expected value exceeds p_T. This is precisely (36).

Suppose we have proved the result for any $t' > t$. Going backwards to period t, there are two cases to consider:

Case 1: (36) holds. Let $v_t = \mathbb{E}[v \mid s_1, \ldots, s_t]$, then (36) becomes

$$v_t - p_t > \delta^{\tau-t} \left(v_t - p_\tau \right), \quad \forall \tau \in \{t+1, \ldots, T, \infty\},$$

In this case we want to show that the buyer should purchase in period t. Indeed, if she does not purchase, then she thinks it is possible for nature to provide no information in the future. Under this particular information process, her future self will purchase in some period $t' \in \{t+1, \ldots, T, \infty\}$, where $t' = \infty$ indicates none of her future selves purchases.[44] Thus from period t buyer's perspective, her expected payoff from continuing into the future is $\delta^{t'-t}(v_t - p_{t'})$ under this process, which by assumption is smaller than $v_t - p_t$. Hence the *worst-case* payoff from delaying purchase is worse than the payoff from purchasing in period t.

Case 1: (36) does not hold. We continue to use v_t to denote the buyer's expected value in period t, then in this case we have the opposite inequality

$$v_t - p_t \leq \max_{\tau > t} \delta^{\tau-t} \left(v_t - p_\tau \right).$$

---

[44]Given our induction hypothesis, this t' is the first period after t such that $v_t - p_{t'} > \delta^{\tau-t}(v_t - p_\tau), \quad \forall \tau \in \{t'+1, \ldots, T, \infty\}$. But this is not important for our argument.

We will show that the buyer should *not* purchase in period t. Specifically, we show that from period t buyer's perspective, her expected payoff from delaying purchase is at least $v_t - p_t$ under *any* future information process, so long as her future selves will act according to induction hypothesis. For this, we claim a stronger result that given the behavior of future selves, expected payoff from delaying purchase is at least $\max_{\tau>t} \delta^{\tau-t}(v_t - p_\tau)$, under *any* future information process.

This stronger claim can be proved by another backward induction. Suppose it holds for $t+1$. In the case of period t, we fix an information process that starts in period $t+1$, and define v_{t+1} to be the buyer's expected value in period $t+1$, under this process. If the signal in period $t+1$ is sufficiently positive that $v_{t+1} - p_{t+1} > \delta^{\tau-t-1}(v_{t+1} - p_\tau)$ for all $\tau > t+1$, then the period $t+1$ buyer optimally purchases in that period. Expected payoff discounted to period $t+1$ is thus $v_{t+1} - p_{t+1} = \max_{\tau>t} \delta^{\tau-t-1}(v_{t+1} - p_\tau)$. Otherwise, the buyer delays purchase, which also yields expected payoff at least $\max_{\tau>t+1} \delta^{\tau-t-1}(v_{t+1} - p_\tau) = \max_{\tau>t} \delta^{\tau-t-1}(v_{t+1} - p_\tau)$, by the induction hypothesis.

Hence, under any information process, expected payoff discounted to period $t+1$ is at least $\max_{\tau>t} \delta^{\tau-t-1}(v_{t+1} - p_\tau)$. It follows that expected payoff from period t perspective is at least

$$\delta \cdot \mathbb{E}\left[\max_{\tau>t} \delta^{\tau-t-1}(v_{t+1} - p_\tau)\right] \geq \delta \cdot \max_{\tau>t} \mathbb{E}\left[\delta^{\tau-t-1}(v_{t+1} - p_\tau)\right] = \max_{\tau>t} \delta^{\tau-t}(v_t - p_\tau).$$

This proves the above claim as well as the whole lemma. $\qquad\square$

We can use this lemma to show that our main result is unchanged in this model with buyer uncertainty (about future information).

Proposition 8. *Under the same assumption as the preceding lemma, a constant price path of p^* achieves the seller's optimal profit guarantee of Π^*.*

Proof. On one hand, with a constant price path of p^*, the seller ensures that the uncertainty-averse buyer only purchases in period 1 (because she anticipates no information in any future period). Thus, by our one-period analysis, the resulting profit is at least Π^*.[45]

On the other hand, we can use the same construction as in Lemma 2 to show that the seller cannot get higher profit in the worst case. Specifically, recall the information process constructed in the proof of Lemma 2. For the buyer who is told her value is below the current price-dependent threshold $G(w_{t_j})$, her expected value makes her indifferent between purchasing now and continuing without further information. Since no future information is the worst case with buyer uncertainty, it is optimal for such a buyer to delay purchase just as in our main model. Similarly, for the buyer whose value is above $G(w_{t_j})$, she should purchase in the current period regardless of

---

[45]Note that buyer uncertainty simplifies this part of the argument, which used to require the Replacement Lemma in our main model.

her true value, which remains true even with uncertainty. Thus under this information process, the distribution of purchase times of an uncertainty-averse buyer is no different from our main model. Hence the seller's profit is also the same, which is bounded above by Π^* as we have shown. $\qquad\square$

F. OTHER RESULTS

F.1. Known Information Arrival Process

This appendix walks through details of the example in Section 1.1. Suppose the prior F is such that $\mathbb{P}[v = 4] = \frac{1}{4}$, $\mathbb{P}[v = 3] = \frac{1}{2}$ and $\mathbb{P}[v = 0] = \frac{1}{4}$.

Case 1: Buyer learns whether or not $v = 4$ in period 1, and learns v in period 2. A buyer who learns $v \neq 4$ has expected value for the object equal to 2. So in order to sell to such a buyer in period 1, the seller's price is at most 2 in the first period. The highest profit under such a selling strategy is 2.

Suppose instead that the seller sets prices so that a buyer with $v \neq 4$ does not purchase in period 1. Then, either the seller gives up selling to such buyers altogether, or his second period price is at most 3 in order to sell to a buyer with $v = 3$. In the former case, profit comes only from the buyer with $v = 4$ and is thus bounded above by $\frac{1}{4} \cdot 4 = 1$. In the latter case, profit from the buyer with $v = 4$ is bounded above by $\frac{1}{4} \cdot (4 - \delta)$, where $4 - \delta$ is the highest price that can be charged in period 1 such that this buyer does not delay (since delaying gives expected payoff $\delta \cdot (4 - p_2) \geq \delta$). Combined with the observation that profit from the buyer with $v = 3$ is at most $\delta \cdot \frac{1}{2} \cdot 3 = \frac{3\delta}{2}$, we conclude that total profit under this strategy is at most $\frac{1}{4} \cdot (4 - \delta) + \frac{3}{2}\delta = 1 + \frac{5}{4}\delta$.

The optimal profit is thus $\max\{2, 1 + \frac{5}{4}\delta\}$, and which strategy ($p_1 = 2, p_2 \geq 3$ versus $p_1 = 4 - \delta, p_2 = 3$) is better depends precisely on whether or not $\delta \leq \frac{4}{5}$.

Case 2: Buyer learns whether or not $v = 3$ in period 1, and learns v in period 2. In this case, the seller can again obtain profit 2 by selling to everyone at a price of 2 in the first period (since a buyer who knows $v \neq 3$ has expected value 2).

Alternatively, the seller does not sell to a buyer with $v \neq 3$ in the first period. In this case the prices $p_1 = 3, p_2 = 4$ leave the buyer with no surplus, and generate maximal profit $\frac{1}{2} \cdot 3 + \delta \cdot \frac{1}{4} \cdot 4 = \frac{3}{2} + \delta$. This latter strategy is better precisely when $\delta > \frac{1}{2}$, as described in the Introduction.

Maxmin optimal price and profit. Next, we compute the G transformation from the distribution F, where $F(v) = \frac{1}{4}$ for $v \in [0, 3)$, $F(v) = \frac{3}{4}$ for $v \in [3, 4)$, and $F(v) = 1$ for $v \geq 4$. Following Definition 1' at the beginning of Appendix A, we compute:

$$g(\alpha) = \begin{cases} 0 & \alpha \leq \frac{1}{4} \\ \frac{3(\alpha - \frac{1}{4})}{\alpha} & \frac{1}{4} < \alpha \leq \frac{3}{4} \\ \frac{3(\frac{1}{2}) + 4(\alpha - \frac{3}{4})}{\alpha} & \frac{3}{4} \leq \alpha \leq 1 \end{cases} \quad .$$

The inverse function of $g(\alpha)$ gives us the pressed distribution G:

$$G(p) = \begin{cases} 0 & p < 0 \\ \frac{3}{4(3-p)} & 0 \leq p < 2 \\ \frac{3}{2(4-p)} & 2 \leq p < 2.5 \\ 1 & p \geq 2.5 \end{cases}$$

One can verify that $p(1 - G(p))$ is decreasing for $p \in [2, 2.5]$. Hence we have:

$$p^* = \underset{0 \leq p \leq 2}{\arg\max} \, p \left(1 - \frac{3}{4(3-p)} \right) \Rightarrow 1 - \frac{3}{4(3-p^*)} - \frac{3p^*}{4(3-p^*)^2} = 0 \Rightarrow p^* = \frac{3}{2}.$$

Verifying that the objective function is concave on $[0, 2]$ means that we have indeed found a global maximum. The profit guarantee corresponding to this price is $\Pi^* = \frac{3}{4}$.

Finally, at the price $p^* = \frac{3}{2}$, the worst-case information structure recommends purchase with (conditional) probability 1 when $v = 4$, with probability $\frac{1}{2}$ when $v = 3$, and with probability 0 when $v = 0$.

F.2. Alternative Interpretation of Π^*

In this appendix, we consider a game where the buyer (rather than nature) chooses information, but where Π^* also emerges as the seller's equilibrium profit. 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 (single) buyer first chooses an information structure $\mathcal{I} : \mathbb{R}_+ \to \Delta(S)$. The seller then chooses a price $p \in \mathbb{R}$ to maximize his profit. 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 the distribution of posterior expected values F_W^B (as described in Section 5.1). This turns out to simultaneously minimize the seller's profit.

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 depending on the price. That is, we maintain the same setup as in Roesler and Szentes (2017), except that the buyer chooses a price-dependent information structure $\mathcal{I} : V \times P \to \Delta(S)$.[46]

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

Proposition 9. *Consider the above information acquisition game where the buyer chooses a price-dependent information structure. In any buyer-optimal Nash equilibrium of this game, the seller's profit is Π^* and the buyer's expected payoff is $\mathbb{E}[v] - \Pi^*$.[47]*

Similar to Roesler and Szentes (2017), trade occurs with probability 1 in equilibrium. However, the buyer's optimal payoff is higher in this game than in that model.

Proof of Proposition 9. For each price p, let $\mathcal{I}^*(p)$ be the corresponding worst-case threshold information structure in our main model. We first construct a Nash equilibrium as follows: 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 $p = \Pi^*$ against *this* price-dependent information structure, and $p = p^*$ against any other information structure (which is off the equilibrium path).

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 the seller's profit is Π^* and the 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. It remains to check that the buyer cannot profitably deviate. Indeed, by setting the price to be p^*, the seller obtains profit at least Π^* when the buyer deviates. Since total surplus cannot exceed $\mathbb{E}[v]$, buyer's payoff is at most $\mathbb{E}[v] - \Pi^*$. This verifies our equilibrium construction.

From the above discussion, we also know that the seller can always choose $p = p^*$ and guarantee Π^*. So in *every* equilibrium, seller profit is at least Π^* and buyer payoff cannot exceed $\mathbb{E}[v] - \Pi^*$. This proves that the constructed equilibrium is buyer-optimal. $\square$

Note that the same argument works for an arbitrary horizon. That is, suppose the buyer chooses a (price-dependent) dynamic information structure to maximize her payoff, whereas the seller responds with a pricing strategy. Then in every buyer-optimal equilibrium of this game, the buyer receives $\mathbb{E}[v] - \Pi^*$ and the seller obtains Π^*.

---

[46]We implicitly require the buyer to commit to acquiring information according to $\mathcal{I}$ after the price is realized. A different interpretation is that such information may be provided by a third party whose objective is to help the buyer (rather than directly hurt the seller).

[47]We thank an anonymous referee for pointing out that sub-game perfect Nash equilibrium may not exist, because there are price-dependent information structures against which the seller has no best response.

F.3. Uniqueness of Du's Mechanism

Recall the random price mechanism from Section 5 and further discussed in Online Appendix B.1. In general, there could be more than one point S for which (24) holds. If that was the case, the seller's optimal strategy in the one-period model with price-independent information would not be unique.

Nonetheless, the point S is indeed unique for *generic* distributions F.[48] The intuition is simple: (24) must bind at some S when W is smallest possible (subject to F being a mean-preserving spread of F_W^B). But for (24) to bind at two different points S would impose a non-generic constraint on F. We omit the formal proof of this genericity result, which is tangential to the paper.

In the following result, we verify that the optimal price distribution is unique whenever S is uniquely defined.

Lemma 10. *There is a uniquely-optimal random price distribution in the one-period price-independent model if and only if (24) holds at a unique point S.*

Proof. "Only if" follows from Online Appendix B.1, so we focus here on the "if" direction. Suppose S is unique, we need to show any random price that guarantees W must be distributed according to $D(\cdot)$. Let $h(p)$ be the p.d.f. of the random price, then seller's profit is given by

$$\Pi = \int_0^1 p \cdot h(p) \cdot (1 - \tilde{F}(p)) \, dp. \tag{37}$$

where $\tilde{F}$ represents the distribution of posterior expected values that nature chooses to minimize Π. Nature's constraint is that F 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), choosing $\tilde{F} = F_W^B$ forces $\Pi \leq W$. On the other hand, seller's optimal pricing strategy guarantees $\Pi \geq W$. So W is the value of the zero-sum game between seller and nature, and whenever the seller uses an optimal strategy, $\tilde{F} = F_W^B$ is a solution to nature's problem. By assumption, the above integral inequality constraint only binds at $x = S$ when $\tilde{F} = F_W^B$. Standard perturbation techniques thus imply that $\tilde{F} = F_W^B$ is nature's optimal choice only if $p \cdot h(p)$ is a constant for $p \in (W, S)$. Indeed, suppose that $p \cdot h(p) > p' \cdot h(p')$

---

[48]A sufficient condition for S to be unique is that $xF(x)$ is strictly convex. To see this, note that $x(F(x) - F_W^B(x)) = xF(x) + W - x$ is strictly convex, so it has at most two roots $x_0 < x_1$. Since $F(x) > F_W^B(x)$ for $x < x_0$, (24) implies S cannot be the smaller root x_0. Hence S must be the bigger root x_1.

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 is still feasible, but the profit is reduced. Similarly, $p \cdot h(p)$ must also be a constant on the interval $p \in (S, B)$. Let c_1, c_2 be these constants.

We now show $c_2 = 0$. Observe that $h(p)$ must be supported on $[W, B]$. So we can alternatively write

$$\Pi = c_1 \int_W^S (1 - \tilde{F}(p)) \, dp + c_2 \int_S^B (1 - \tilde{F}(p)) \, dp.$$

Let nature fix $\tilde{F}(p) = F_W^B(p)$ for $0 \leq p \leq S$. Then $\int_S^1 (1 - \tilde{F}(p)) \, dp = \int_S^1 (1 - F_W^B(p)) \, dp = \int_S^1 (1 - F(p)) \, dp$. This yields

$$\Pi = c_1 \int_W^S (1 - F_W^B(p)) \, dp + c_2 \int_S^1 (1 - F_W^B(p)) \, dp - c_2 \int_B^1 (1 - \tilde{F}(p)) \, dp.$$

Given the seller's choice of c_1, c_2, the first two terms above are constants. So nature's problem is to choose $\tilde{F}(p)$ for $p \in (S, 1)$ to maximize $c_2 \int_B^1 (1 - \tilde{F}(p))$. Since $\int_B^1 (1 - F_W^B(p)) = 0$, $\tilde{F} = F_W^B$ can only be an optimal choice when $c_2 = 0$.

To summarize, we have shown that the seller's price density $h(p)$ must be supported on $[W, S]$ and $p \cdot h(p)$ is a constant. This condition together with $\int_W^S h(p) \, dp = 1$ uniquely pins down $h(p)$, which is exactly the density function of $D(x)$. Lemma 10 follows. $\qquad\square$

F.4. 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:

Lemma 11. $\Pi_{RSD} \geq \Pi^*$ *with equality if and only if* $W = \underline{v} = p^*$. *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 $W = \Pi_{RSD} = \Pi^*$, 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, the seller's profit is either $W = \Pi^*$ (when $p^* < B$) or 0. Since we showed in our main model that the seller can guarantee Π^* with a price of p^*, profit must be W and the Roesler-Szentes information structure is a worst case for the price p^*. Thus $W \geq p^*$, because a worst-case information structure cannot induce a posterior expected value strictly below p^*. We there conclude $p^* = W = \Pi^* = p^*(1 - G(p^*))$, from which it follows that $G(p^*) = 0$ and $p^* = \underline{v}$. Thus $W = \underline{v}$ must hold.

Finally, the ratio Π_{RSD}/Π^* is unbounded even within distributions F that have binary support. This follows from Proposition 6 in Carrasco et al. (2018). However, we conjecture that this profit ratio becomes bounded under certain regularity conditions on F. □