Informational Robustness in Intertemporal Pricing

Jonathan Libgober* and Xiaosheng Mu**

*University of Southern California
**Cowles Foundation and Columbia University

January 17, 2019

Abstract. We introduce a robust approach to the study of optimal dynamic pricing in the face of information arrival. As opposed to classical intertemporal pricing models, we consider consumers who learn about their willingness-to-pay for a product over time (which is typical for initially unfamiliar products). A seller commits to a pricing strategy, while buyers arrive exogenously and decide when to make a one-time purchase. The seller does not know how each buyer learns about her value for the product, and seeks to maximize profits against the worst-case information arrival processes. We show that a constant price path delivers the robustly optimal profit, 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, xm2230@columbia.edu. We are particularly indebted to Drew Fudenberg for guidance and encouragement. We also thank Ben Brooks, Gabriel Carroll, Rahul Deb, Dirk Bergemann, Yeon-Koo Che, Songzi Du, Kfir Eliaz, Ben Golub, Jerry Green, Johannes Hörner, Yuhta Ishii, Navin Kartik, Scott Kominers, Eric Maskin, Tomasz Strzalecki, Juuso Toikka, and seminar audiences at Australian National University, ASU, Columbia, Cornell (Johnson), Harvard, Illinois, LA Theory Day, Michigan, MIT, Monash, NYU, Rochester, Stony Brook, University of Technology-Sydney, UNSW, USC, and Yale for comments. Any remaining errors are ours.

1. Introduction

1.1. Motivation and Results

Suppose a monopolist has invented a new durable product, and is deciding how to set prices over time to maximize profit. Since the product is new, consumers could become more familiar with it as they learn more about it. For example, when the Apple Watch, Amazon Echo, and Google Glass were released, most consumers had little prior experience to inform their willingness-to-pay. In such a situation, the monopolist might suspect that purchase decisions will depend on the available information–e.g., journalist reviews about the product–which may in turn depend on pricing. The possibility of information arrival and its potential to delay consumer purchase present challenges to the monopolist's problem.

Some prior work on intertemporal pricing has studied how fluctuations in buyer willingness-to-pay can influence a monopolist's optimal selling strategy—see Deb (2014) and Pavan, Segal and Toikka (2014) among others. As we discuss in Section 1.2, this literature has traditionally assumed that the seller knows exactly how a buyer's willingness-to-pay evolves. Under this approach, a general lesson is that the optimal selling strategy should leverage the evolution of buyer (expected) values. But depending on the environment, different conclusions emerge regarding whether information arrival would favor increasing or decreasing price paths. In this paper, we differ by considering the case where the seller may not exactly know how buyers receive information, which is especially relevant for new products.

In our model, buyers arrive exogenously and observe signals of their values over time, each according to some information structure (or more precisely, information arrival process). A feature of our model is that we allow information to depend on past and current prices—for example, if consumers see more about products that are cheaper. However, this information structure is *unknown to the seller.* With limited knowledge of how buyers learn, the seller commits to a pricing strategy to *maximize his profit guarantee* against all possible information structures. For instance, the seller may be aware that there are a large number of product review websites. But if the information consumers seek relates to their idiosyncratic needs, the seller may not know which buyers have access to which websites (and when). He may therefore 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 this price path, buyers do not delay purchase in the *worst-case* information structure. The striking feature is that the intertemporal incentives of buyers, suggested by the prior literature to be an important determinant of optimal pricing, do not matter if the seller adopts an informationally robust objective. We note that constant price paths are known to be optimal in settings without buyer learning, as in Stokey (1979) or Conlisk, Gerstner and Sobel (1984). But their results rely upon the observation that buyers with known values never gain from purchasing later at the same price. In our setting, however, future information arrival can encourage delay even against a constant price path. Thus a key part of our analysis is to show that such delay would only improve the seller's profit. The rough intuition is that when prices are held constant, a buyer is only willing to delay if he expects (at least) the same discounted probability of purchase in the future. This necessarily translates into (weakly) higher profit for the seller from delayed sale. We also emphasize that the optimal price itself differs from the known value case. So while intertemporal incentives play no role, the possibility of information arrival does influence the price.

Why does the seller not benefit from intertemporal price discrimination? We first recall the classic intuition from the known value case: Although lowering prices over time increases the sale to low-value buyers, it also causes some high-value buyers to delay purchase, leading to a negative net effect on profit. This argument does not immediately extend to our setting, since the net effect on profit may be positive for a *fixed* information arrival process.[1] However, it does generalize under the robust objective, meaning that *worst-case* information always entails greater loss from delayed purchase than the benefit from increased sale. We show this by focusing on the class of *partitional information arrival processes*, which informs the buyer in each period whether her value is above or below a threshold. We consider thresholds so that a buyer with value below the current threshold is indifferent between buying now or waiting. This particular information structure makes the seller's problem separable across time, eliminating potential gains from intertemporal price discrimination.

Our analysis echoes others in the robust mechanism design literature, which highlight that simple strategies can be optimal given sufficient uncertainty over the environment. Constant price paths are "simple" because the optimum can be achieved without knowing the buyers' arrival times or even the time horizon. Obtaining a result of this form in our setting provides justification for firms to eschew sophisticated pricing strategies, even when consumer learning is significant. We find it reassuring that the worst-case information structure is always partitional (as introduced above). Partitional information structures have been studied in settings without the robust approach, as in Kolotilin (2015).[2] So long as the seller seeks robustness against this restricted class of processes, our analysis is unaffected.

The constant price path result serves as a benchmark that can be used to explain how features of the seller's problem translate into certain pricing patterns. While the possibility of buyer

[1] See Appendix D.2 for an example.

[2] Also related are Bergemann and Wambach (2015) and Li and Shi (2017). These papers discuss that partitional information arrival may arise via comparisons to past products for which buyer values are known. In the case of the iPhone, for instance, consumers may compare the most recent iteration to past cell phone purchases. Partitional processes may also arise if the product has several attributes revealed to the buyer sequentially, with lexicographic preferences over these attributes.

learning does not by itself favor price dynamics, economically meaningful restrictions on the learning environment may cause decreasing or increasing price paths to become optimal. We explore several extensions of our main model in this spirit: First, we consider cases where buyers receive information infrequently. We show that a declining price path out-performs a constant (or increasing) price path. Second, we study a situation where buyers have access to at least some initial information. Price dynamics can arise if this initial information differs across buyers who arrive at different times. Finally, we present a variant of our model with common values and public signals. We characterize the optimal pricing strategy in the patient limit, which involves prices that increase over time. This is true even though information in our model is not generated by previous sales, so our result offers a new justification for introductory pricing. In light of these possible departures from constant pricing, our main result says that a constant policy is nonetheless optimal in the presence of very rich informational uncertainty.

Relative to the existing literature on robust (static) mechanism design, dynamic information arrival presents certain conceptual modeling challenges. One such challenge we highlight in the paper is that when information is revealed over time, there are potentially many ways to model the interaction between information and prices. Our main model considers the case where information in each period can depend on prices up to and including that period. We view this generality as desirable, since in practice prices can influence information availability (see further discussion in Section 2.3). An alternative would have been to disallow such price-dependence, as studied by Du (2018) in a one-period model, building on the earlier work of Roesler and Szentes (2017). In Section 5, we extend their analysis to the dynamic setting and resolve whether constant price paths can still deliver the robustly optimal profit without price-dependent information.

We briefly discuss the technical innovations that underly our results, as they may be applicable to related problems, particularly those that involve dynamic 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 *given a pricing strategy*. In particular, our Proposition 2 shows it is always without loss to restrict attention to partitional information structures. 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 (2018).

Another useful tool in our analysis is the Replacement Lemma, stated variously as Lemma 1, Lemma 3 and Lemma 4 for different variations of our model. This set of lemmata gives sufficient conditions on prices to guarantee that the worst-case information structure does not involve dynamics. Our technique is as follows: By modifying the timing and probability of nature's "recommendation to purchase," we can replace an arbitrary dynamic information structure with a static one while reducing the seller's profit. When such a result obtains, buyers do not delay purchase in the worst case, and our dynamic analysis simplifies to a static problem. The method of replacement we develop here is not limited to the worst-case objective and may be applied more broadly.

Below we first review the literature, and then proceed to present the main model. Section 3 analyzes the one-period model, and we show the optimality of constant price paths in Section 4. Section 5 discusses our assumption regarding price-dependent information, while Section 6 presents other extensions of our model. Section 7 concludes.

1.2. Related Literature

This paper is part of an active literature that studies pricing under robustness concerns, where the seller may be unsure of some parameters of the buyer's problem. Informational robustness is a special case, and one that has been studied in static settings. The most similar to our one-period model is the single-buyer case in Du (2018). He considers a setting like ours, where the buyer's value comes from some commonly known distribution, but the seller does not know the information structure that informs the buyer of her value. However, Du (2018) assumes that nature's choice of information structure is independent of the price. The one-period version of our model differs by assuming that the seller first posts a price, and then nature can reveal information depending on the *realized* price. This timing difference (regarding whether nature moves after, or simultaneous to, the seller) is further discussed in Section 2.3. By studying a dynamic model, our focus is on robustness of the seller's pricing strategy to potential buyer delay, which is absent from Du (2018).

The worst-case information structure in Du (2018) first appears in Roesler and Szentes (2017), who study a related model where the buyer chooses her optimal information structure, after which the seller posts a price to maximize profit (without facing uncertainty). Roesler and Szentes (2017) show that the solution also minimizes the seller's profit. In Appendix D.3, we provide a related Bayesian interpretation of our model and results. See also Section 5.2 for the role of the Roesler-Szentes information structure in a dynamic generalization of Du (2018)'s model.

Other papers have studied the case where the value distribution itself is unknown to the seller.[3] For instance, Carrasco et al. (2018) consider a seller who does not know the distribution of the buyer's value, but who may know some of its moments.[4] We note that knowing the mean and the range of the value distribution is equivalent to our model with a prior distribution having two-point support. In this sense, informational uncertainty nests value uncertainty when

[3]Related to Roesler and Szentes (2017), Condorelli and Szentes (2018) study the problem where the buyer chooses her optimal value distribution.

[4]Randomized pricing is optimal in Carrasco et al. (2018) precisely because their timing assumption coincides with Du (2018), as described in the previous paragraph.

only the first moment is considered. But in general, even in the static setting, assuming a prior distribution constrains the possible posterior distributions nature can induce beyond any set of moment conditions.[5] Prior literature has also studied pricing under value uncertainty with different non-Bayesian objectives, such as minimax regret—see Bergemann and Schlag (2011), Handel and Misra (2014), Caldentey, Liu and Lobel (2016), Liu (2016) and Chen and Farias (2018).

Our use of the informationally robust objective is inspired in part by Bergemann, Brooks and Morris (2017), Brooks and Du (2018) and Du (2018). The goal of this line of research is to move away from specific assumptions about the informational environment, which (as we discuss below) may imply optimal mechanisms that depend sensitively on these assumptions. Relative to these existing work, we introduce *dynamic* informational robustness and illustrate conceptual issues that arise in formalizing this notion.[6] Within the broader literature on robust mechanism design, our constant price path result fits the general agenda of providing optimality foundations for certain simple mechanisms observed in practice. For instance, Carroll (2017) shows how uncertainty over the correlation between a buyer's demand for different goods leads the seller to price the goods independently.[7] A similar theme runs through many other papers as well; see Chung and Ely (2007), Frankel (2014), Carroll (2015) and Yamashita (2015).

As mentioned above, the existing literature on intertemporal pricing suggests that changes in buyer willingness-to-pay may lead to gains to non-constant pricing policies. Stokey (1979) shows that the optimal price decreases over time if high-value consumers lose more from waiting. This complements another (perhaps better-known) result in Stokey (1979) that constant prices would be optimal if all consumers were to know their values and discount equally, a result which we generalize in Proposition 4. Deb (2014) assumes the value is independently redrawn upon Poisson shocks and obtains increasing prices as optimal.[8] Garrett (2016) finds cyclical pricing to be optimal in a model with arriving buyers whose values follow a two-type Markov-switching process. More generally, recent work on dynamic mechanism design (Courty and Li (2000); Pavan,

[5]The working paper version of Carrasco et al. (2018) contains an extension to multiple periods and repeated sales. While they also find repetition of the static pricing rule to be optimal, this is because in their setting buyer demand is *reset* every period. In contrast, we focus on the case of durable goods, so that buyers face intertemporal tradeoffs. As we discuss below, intertemporal incentives by themselves may lead to gains to non-constant pricing. Thus our result is distinguished from Carrasco et al. (2018).

[6]As far as we are aware, Chassang (2013) and Penta (2015) are among the few papers that study a dynamic robust objective, but these are both rather different from our setting. Penta (2015) considers the dynamic implementation of social choice functions, and Chassang (2013) shows how dynamics enable a principal to approximate robust contracts that may be infeasible under liability constraints.

[7]The general link between dynamic allocations and multi-dimensional screening has been noted in Bayesian mechanism design settings (see e.g. Pavan, Segal and Toikka (2014)). 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.

[8]The earlier work of Conlisk (1984) considers a two-period model where buyer values have binary support and are redrawn in the second period. The result there is that decreasing prices are optimal for some parameters, in contrast to Deb (2014).

Segal and Toikka (2014)) has explicitly characterized optimal selling mechanisms when the seller knows how buyer values evolve. The solution often depends sensitively on assumptions regarding how buyers learn about their values (i.e., how their expected values evolve). We are primarily interested in the case where the seller faces uncertainty over how buyers receive information, and consequently our findings are in contrast to these results.

In relation to the Bayesian persuasion literature (Kamenica and Gentzkow (2011), Ely (2017)), we also allow general information structures to inform the buyers of their values. But we differ from persuasion models by looking at the strategic interaction between information and pricing. We develop new tools such as the Replacement Lemma to characterize worst-case information processes for forward-looking buyers. We then use this characterization as an intermediate step toward solving the optimal pricing strategy.

2. Model

Our baseline model adds buyer learning to an otherwise straightforward dynamic pricing setting. Justification of our modeling choices can be found in Section 2.3. A seller (he) sells a durable good at times $t = 1, 2, \ldots, T$, where $T \leq \infty$. In each period t, a single buyer (she) arrives.[9] We let t denote calendar time, and let a index a buyer's arrival time. All parties have discount factor δ. The product is costless for the seller to produce,[10] while each buyer has unit demand. We assume that each buyer has (undiscounted) lifetime value v_a from purchasing the object, where v_a is drawn from a distribution F and fixed over time; when there is no confusion, we will omit the subscript and simply write the value as v. The prior distribution F is common knowledge, with support on $\mathbb{R}_+$ and $0 < \mathbb{E}[v] < \infty$. Until Section 6.3, we assume different buyers have independent values.

At time 0, the seller commits to a pricing strategy σ, which is a distribution over possible price paths $p^T = (p_t)_{t=1}^T \in \mathbb{R}_+^T$. Note that the price the seller posts at time t must be the same for all buyers that have arrived and not yet purchased (see Section 2.3 for discussion). When a buyer arrives, she 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.* The next subsection formalizes the learning process. A buyer who never purchases the object obtains a payoff of 0.

[9]Buyer arrivals have been studied elsewhere in the dynamic pricing literature; see Conlisk, Gerstner and Sobel (1984), Board (2008) and Garrett (2016). Our analysis is unchanged if the number of arriving buyers varies over time, so long as the value distribution is fixed.

[10]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.

2.1. Dynamic Information Structures

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

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

Without loss of generality, we assume that all buyers are endowed with the same signal sets S_t, although each one privately observes any particular signal realization. To be fully rigorous, there should be a σ-field associated with each S_t, and the mappings $I_{a,t}$ 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 $v_a \in \mathbb{R}_+$, the history of her previous signal realizations $s_a^{t-1} = (s_a, s_{a+1}, \ldots, s_{t-1}) \in S_a^{t-1}$, as well as the history of *all previous and current prices* $p^t = (p_1, p_2, \ldots, p_t) \in P^t$. The possibility for information to flexibly depend on realized prices distinguishes our model from most existing literature, 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 a standard definition of dynamic information structures. The signal s_t would occur with probability $I_{a,t}(s_t \mid v_a, s_a^{t-1})$, where we can omit the dependence on $(p_t)_{t=1}^T$ since it is the only possible realization.[11] Allowing for price-dependent information only has bite when the seller randomizes.[12]

2.2. Seller's Objective

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

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

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

[11] Even though in this case information revelation will not depend on the particular price path, the *worst-case* information structure generally will.

[12] 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.

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

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

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

Note that when a buyer faces indifference, ties are broken against the seller. It will follow from our analysis that $\sup\inf$ is achieved as $\max\min$. 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 of Assumptions

In this subsection we comment on several of our key modeling assumptions.

Bayesian buyers and uncertain seller. We assume that each buyer knows her information arrival process, and is Bayesian about what information will be received in the future. The seller, on the other hand, knows only that there are many possible ways buyers can learn. In line with the robust mechanism literature, we find it reasonable that buyers are much better informed than the seller about the informational environment. In our setting, for instance, who tends to watch which commercials or visit which product review websites is often idiosyncratic and beyond the seller's knowledge.

It may seem unrealistic that buyers perfectly know the signal distribution far into the future. However, we mention that due to the stopping time nature of buyer decision, our analysis is unchanged if buyers also face uncertainty (and are maxmin) over future information, so long as they can interpret signals in the current period.[13] In this sense, our results do not rely on extra rationality of the buyers beyond what is typically assumed in *static* robust mechanism design.

Although our main model assumes extreme uncertainty faced by the seller, Section 6.2 considers an extension where the seller believes that buyers have *at least* some information (on top of the common prior F). A constant price path (with a potentially different price) remains optimal, strictly generalizing the qualitative finding of Stokey (1979) in the known value setting.

Price-dependent information. Our key assumption in defining dynamic information structures is that information in each period can depend on realized prices up to and including that period. In reality, such price-dependence occurs through a number of channels: Ordered display on product websites, selective coverage by reviewers and rational inattention of buyers. Thus we

[13] Formally developing that extension requires a different modeling framework, so we omit the details.

view the above setup as the most natural specification of our model. Allowing nature to provide price-dependent information also delivers the most robust profit guarantee for the seller.

In Section 5, we consider a variant of our model without price-dependence. In one period, that model reduces to the static problem studied in Du (2018). We solve a dynamic version of his single-buyer problem and show that a randomization over constant price paths is optimal. We also discuss the case of limited price dependence.

On the conceptual level, 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.* This point is further discussed and formalized in the sequential decision theory framework in Ke and Zhang (2018). In our explicitly dynamic setting, the timing issue is more salient because there are many more ways one could model the timing of nature's moves. For example, our main model takes the most pessimistic perspective that nature moves in each period, after the seller's randomization. A contribution of this paper is to delineate various timing assumptions and highlight their implications.

Private signals. Other than price-dependence, our main model assumes that buyers receive information privately. This is without loss when buyer values are independent, in which case profit can (at worst) be minimized on a per buyer basis.[14] However, if buyer values were correlated, requiring information to be publicly released would limit nature's ability to damage profits. See Section 6.3 for such a model and further discussion.

Commitment. We believe that many firms are able to credibly announce and stick to consistent pricing strategies, due to reputational concerns. And while some strategies may be difficult for a seller to commit to, constant price paths are simple to implement since deviations are straightforward to detect. Dropping commitment from the model would also introduce technical difficulties related to formalizing learning under ambiguity; see Epstein and Schneider (2007).

More complex pricing strategies. In some markets, the seller may be able to discriminate buyers based on their identity (i.e., arrival time) and charge personalized prices. While personalized pricing may have benefits in general, such strategies turn out to not improve the seller's profit guarantee in our main model.[15] Similarly, pricing strategies that condition on the sales history have no benefits in our model.

That said, we do restrict the seller to using pricing mechanisms, and rule out for instance

[14] To be more specific, in our main model there is no need for nature to correlate information across buyers, or to condition a buyer's signal on the purchase history of other buyers.

[15] To see this, first observe that nature can release information to simultaneously minimize the seller's profit from each buyer. So the total profit guarantee cannot exceed the sum of what can be ensured from individual buyers. (This holds unless buyer values are correlated *and* information is public, which is studied in Section 6.3.) On the other hand, this upper-bound is achieved by a constant price path, which makes the problem stationary across buyers. Thus our main result suggests that the seller need not use personalized pricing, as we assumed above.

mechanisms that randomly allocate the object as a function of each buyer's report. We view this as a restriction on the environment, but one that is natural in our applications of interest. For online and in-store shopping, for instance, buyers typically observe a posted price rather than a general mechanism (see Dilmé and Garrett (2018) for further discussion). This restriction also allows us to avoid difficulties in working with general dynamic mechanisms, where agent types must capture all future information.[16] We note that whether a general mechanism may improve upon posted prices is closely related to whether information can be price-dependent. This connection is discussed in Section 5.

3. One-period Analysis

We start by analyzing the one-period problem (with a single buyer). To solve this problem, we will define a transformed distribution of the prior F. For expositional simplicity, the following definition assumes F is continuous. All of our results in this paper extend to discrete distributions, though the general definition requires additional care and is relegated to Appendix A.

Definition 1. *Given a continuous distribution F, its "pressed version" G is another distribution defined as follows. For $y \in \mathbb{R}_+$, let $L(y) = \mathbb{E}_F[v \mid v \leq y]$ denote the expected value conditional on the value not exceeding y. 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 structure is outcome-equivalent to another that directly recommends one of two actions: To purchase the good or not. Given this simplification, the worse-case information structure must have the following property. As long as the buyer is recommended to 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, subject to the constraint that a buyer who does not buy has fixed expected value (in our case, p), one can show that *partitional information structures* maximize the probability of this recommendation (see e.g. Kolotilin (2015)). In a partitional information structure, the buyer is told whether her value is above or below a certain threshold. By the above definition of G, this threshold must be $F^{-1}(G(p)) = L^{-1}(p)$, making $1 - G(p)$ the probability of sale.

These remarks give us the following proposition:

[16]The technical issues involved are perhaps magnified under the robust approach, since the seller may additionally ask buyers to report their information *structures*. How to formulate and study such mechanisms is an interesting direction for future work.

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

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

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

It is worth comparing the optimization problem (1) to the standard known value model. 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. This analogy will be useful for the subsequent analysis.

The following example illustrates:

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

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 $1/2$, since such a buyer will purchase in any event. Nonetheless, the lowest element of the partition cannot be further refined. That is, a buyer whose value is below the critical threshold will be told only *this in every worst-case information structure.*

In this example, relative to the known value case, the seller charges a lower price and obtains a lower profit under informational uncertainty. In Appendix D.1, we show that these comparative statics hold for general distributions F.

4. Main Results

With multiple periods and arriving buyers, recall that $\frac{1-\delta^T}{1-\delta}$ is the discounted number of arrivals. The main result of this paper is now stated.

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

To interpret this result, recall that the dynamic pricing literature has shown that when values are known, fluctuations in buyer values can lead to corresponding fluctuations in prices.[17] This

[17] See Stokey (1979), Deb (2014) and Garret (2016), as discussed in the introduction.

insight applies to our setting with buyer learning, since information arrival causes the buyer's expected value for the product to fluctuate. In fact, for a *fixed* information arrival process that is known to the seller, optimal prices can increase or decrease over time, depending on how one specifies the learning process (see Appendix D.2). The lesson is that the seller would in general want to adapt his pricing strategy to how buyer values evolve, *if he knew this evolution.*

In contrast, Theorem 1 suggests that when facing uncertainty over how buyers learn and adopting a robust objective, the seller is best off committing to the simple strategy of a constant price. The underlying mechanism for our result is more involved than the case of known (and constant) values, where the optimality of constant prices follows from Stokey (1979). Indeed, information arrival *may* cause buyers 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 "partitional information structures."

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.[18] 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. 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 can guarantee (undiscounted profit) Π^* from each buyer regardless of nature's choice of information structures. This follows from our Replacement Lemma, which shows that 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^* per buyer. This part of the argument takes advantage of the intuition from the one-period analysis and generalizes the partitional information structure appropriately to the dynamic setting. Below we provide some details of these two parts of our proof, respectively.

[18]We note that Theorem 1 is in itself a separate result from Stokey (1979), as our main model is disjoint from the known value setting. However, in Section 6.2 we present an extension that embeds our main model as well as known values. A constant price path remains optimal in that extension (see the discussion after Proposition 4), thereby generalizing both Stokey's result and Theorem 1.

4.1.1. Lower-bound

Under known values, any 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 non-decreasing price paths, nature cannot hurt the seller more than providing information *only* upon arrival. Applying this lemma to each buyer with prices held constant at p^*, we derive $\Pi^* \cdot \frac{1-\delta^T}{1-\delta}$ as a lower-bound on the seller's total profit guarantee.

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 from the first buyer is minimized by an information structure that only provides information in the first period.*

We call it 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 (from the first buyer). 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 $p_1(1 - G(p_1))$.

To construct such a replacement, we view the original dynamic information structure as providing recommendations to the first buyer to purchase or not at different times. Whenever this buyer was recommended to purchase in period t, in the replacement information structure we have nature recommend her to purchase in period 1 with probability δ^{t-1}. In other words, we "push and discount" nature's recommendation to the buyer's arrival time. 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. Once this is proved, it follows that the discounted probability of sale is unchanged, so that profit can only decrease (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 structures that ensure the seller does not gain more per buyer with a longer horizon. Before getting to the proof, we first introduce a definition of partitional information arrival processes that generalize the partitional information structures in the one-period problem.

Definition 2. *A partitional information arrival process (for the first buyer) involves a descending sequence of (possibly randomized) cutoffs $v_1 \geq v_2 \geq \cdots \geq v_T$, where each v_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 $v \leq v_t$.*

With a single period, we have shown that the worst-case information structure is partitional. The following proposition extends this result to the dynamic setting, suggesting that the seller need only worry about partitional information arrival processes.

Proposition 2 (Worst-case is Partitional)**.** *Given any pricing strategy σ, there exists a partitional information arrival process that minimizes the seller's profit from the first buyer.*

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 partitional information. That said, dynamics does introduce 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 structure with a partitional one, such that the buyer's purchase times are weakly later than before. This stochastic dominance property implies a lower profit.

However, there are two reasons that Proposition 2 is of limited direct use. First, given an arbitrary pricing strategy, it remains challenging to solve for the *exact* worst case (partitional) information process.[19] This is due to difficulties with determining buyer optimal stopping under arbitrary prices and information. Second and perhaps more important, finding the information structure with minimal profit is only helpful insofar as this translates into a profit upper-bound. But the seller's profit is not easily computable in general.

So instead of focusing on the worst case, we appeal to further economic intuitions for our environment to find a particular partitional information process that allows for relatively easy computation of profit. This process will allow us to prove the following lemma:

Lemma 2 (Profit Upper-bound)**.** *For any pricing strategy, there exists a dynamic information structure (for the first buyer) and a corresponding optimal stopping time that lead to profit $\leq \Pi^*$.*[20]

To explain our construction, we assume for simplicity that the seller charges a *deterministic* price path $(p_t)_{t=1}^T$. If the first 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 with value $v \in [w_{t_j}, w_{t_{j-1}}]$ would optimally buy in period t_j. Here w_{t_j} is defined by the indifference condition

[19] The challenge arises primarily when prices decrease over time. With non-decreasing prices, the worst case has been characterized in the Replacement Lemma.

[20] It is crucial for this result that nature can provide information dynamically; see Section 6.1.

$w_{t_j} - p_{t_j} = \delta^{t_{j+1}-t_j} \cdot (w_{t_j} - p_{t_{j+1}})$, 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 dynamic information structure 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 following partitional information arrival process has this property: In each period t_j, the buyer is told whether or not her value is in the lowest $G(w_{t_j})$-percentile, while no information is revealed in other periods. This construction generalizes the one-period analysis, in that thresholds of the partition are chosen to make the buyer *indifferent between purchasing and continuing without further information.* The buyer therefore prefers to delay purchase when her value is below the threshold. On the other hand, a buyer whose value is above the threshold does not expect to receive further information, and hence purchases immediately.

The above observations show that $G(w_{t_{j-1}}) - G(w_{t_j})$ is the probability of sale in period t_j. We can then compute the seller's profit as follows:

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

where we assumed $T = \infty$ for ease of illustration. The second line above is by Abel summation,[21] 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.

When prices are random, the indifference types w_{t_j} will be random variables. A technical difficulty arises because they may not be decreasing over time. When such non-monotonicity occurs, the seller's discounted profit cannot be written as a convex sum of one-period profits, and the profit bound in (2) will not be valid. In Appendix A, we develop additional tools to generalize the above construction. In short, we keep track of the binding indifference types, above which the buyer has already purchased. This focus ensures monotonicity, and we can use the re-defined w_t to specify the partitional information process. The proof of Lemma 2 then extends.

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

To conclude this section, we mention that the above analysis is unchanged if any buyer with value *above* the current threshold perfectly learns her true value, since she will purchase immediately regardless (as in Example 1). In this sense, the partitional information process we construct is outcome-equivalent to one where *higher-value buyers discover their true values earlier*.

We also reiterate that the information structure considered in this proof is generally *not* the worst case (see Example 2 in Appendix A). It is only one process for which buyer's optimal stopping time as well as profit can be tractably determined. Solving for the exact worst-case (partitional) process against a decreasing price path is an open question we hope to address in the future.

5. Price-dependence: Relation to Du (2018)/Roesler-Szentes (2017)

So far in our baseline model, we allow nature to provide information depending on all realized prices. The flexibility in nature's choice delivers the most pessimistic profit guarantee for the seller. In this section, however, we study variants of our model where information does not significantly vary with prices. This alternative assumption is appropriate if, for instance, the seller is widely followed by product reviewers. 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.

Formally, we change the definition of dynamic information structures in Section 2.1 to limit how information depends on realized prices, but we maintain the rest of the setup in Section 2. With a single period, this modification returns the recent model of Du (2018) and connects also to Roesler and Szentes (2017). Below we first review and recast their results in our framework. Moving to the dynamic setting, we then show that when information is *price-independent*, the seller's optimal strategy can be implemented as a randomization over constant price paths. We also comment on the intermediate case where information exhibits limited price-dependence.

5.1. Static Model without Price-dependence

This section describes the modification for our one-period model. To model the alternative assumption that information does not depend on realized price, we assume the distribution over signals is determined by some function $I : \mathbb{R}_+ \to \Delta(S)$, which is price-independent. The seller still has a maxmin objective, but he only worries about this smaller class of information structures.

Our earlier analysis provides some intuition for why randomization strictly helps the seller. To illustrate, suppose $v \sim$ Uniform[0,1] as in Example 1. We have shown that the seller's optimal deterministic price is $p^* = \frac{1}{4}$. Now consider a strategy that randomizes between $\frac{1}{4}$ and $\frac{1}{4} + \varepsilon$, putting a small probability on the latter. We claim this strategy achieves a higher profit guarantee than $\Pi^* = \frac{1}{8}$ against any partitional information structure (that is independent of the realized price). Indeed, if the threshold of the partition is $\frac{1}{2}$ or smaller, the seller sells at both realized prices to buyers with value above the threshold. Profit in this case is seen to be $> \frac{1}{8}$, thanks to the higher price $\frac{1}{4} + \varepsilon$. But if the threshold is larger than $\frac{1}{2}$, then sale would occur with probability one when the price is realized as $\frac{1}{4}$, again leading to profit $> \frac{1}{8}$. A similar but more involved argument shows that such a randomization guarantees profit more than $\frac{1}{8}$ against all price-independent information structures, not just partitional ones.

The above discussion prefigures the results of Du (2018) and Roesler and Szentes (2017), who characterize the optimal price randomization as well as the corresponding worst-case information structure for this one-period model. We proceed to summarize their solutions.

Du (2018) considers a seller who shares our robust objective, but who can use more general mechanisms that prescribe allocation probabilities based on the buyer's reported type. However, it turns out that with a single agent, the same outcome (i.e., profile of interim purchase probabilities) can be implemented using the following *random price mechanism*: The seller charges a random price with c.d.f.

$$D(x) = \begin{cases} 0 & x < W \\ \frac{\log \frac{x}{W}}{\log \frac{S}{W}} & x \in [W, S) \\ 1 & x \geq S \end{cases} \tag{3}$$

for some numbers W and S that depend on the prior distribution F. This mechanism guarantees profit at least W (see Appendix B.1), and it generalizes Proposition 5 in Carrasco et al. (2018), who focus on prior distributions F with binary support.

Roesler and Szentes (2017) consider a related problem where the buyer chooses the information structure to maximize her surplus. The buyer-optimal information structure they identify turns out to be the *minmax information structure* in Du (2018)'s problem. The Roesler-Szentes information structure induces posterior expected values distributed as follows:

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

where W and B are parameters such that F is a mean-preserving spread of F_W^B, and W is smallest possible subject to this constraint. This is the same W that shows up in Du's price distribution.

Taken together, W is the seller's maxmin profit in the one-period problem without price-dependence. 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 Appendix D.5 we characterize when the comparison is strict.

5.2. Dynamic Model without Price-dependence

With multiple periods, we first consider a variant of our main model that completely rules out price-dependent information. Formally, we redefine a dynamic information structure $\mathcal{I}_a$ for the buyer arriving at time a to be

- A sequence of signal sets $(S_t)_{t=a}^T$, and
- Probability distributions given by $I_{a,t} : \mathbb{R}_+ \times S_a^{t-1} \to \Delta(S_t)$, for all t with $a \le t \le T$.

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. Then for any T and δ, maxmin average profit per buyer is Π_{RSD}. The seller can achieve this by randomizing over constant price paths drawn from Du's price distribution $D(x)$ in (3). Nature can force this profit upper-bound by providing the Roesler-Szentes information structure to each buyer upon arrival.*

It is not difficult to understand nature's information choice. By providing the static Roesler-Szentes information structure, nature makes each buyer "know her value" to be drawn from F_W^B. By the result of Stokey (1979), this holds profit below Π_{RSD} from each buyer.

The more striking feature of Theorem 2 is that the seller can guarantee Π_{RSD} via a randomization over constant price paths. We highlight that the presence of randomization allows for many possible ways to *correlate prices across time* while maintaining the same marginal price distribution in each individual period. In this sense, it is notable that the seller's optimal strategy involves extreme correlation in the form of perfectly persistent prices.

To prove that a random constant price delivers the profit guarantee of Π_{RSD}, we demonstrate the following generalization of the Replacement Lemma in Section 4.1.1.

Lemma 3 (Replacement Lemma for Randomized Constant Price Paths)**.** *Suppose that the seller uses a strategy that randomizes over constant price paths, while nature provides information independently of the realized price. Then the seller's profit can be minimized by information structures that only provide information to each buyer upon her arrival.*

This result embeds a special case of the previous Lemma 1 when the seller charges a deterministic constant price. However, the "push and discount" argument we used there to find a replacement 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.

In our proof of Lemma 3, we introduce the concept of "cutoff prices" for a given price-independent information arrival process. These cutoff prices are the dual notion of "cutoff values" used in the proof of Lemma 2, and they capture when the buyer should purchase, and at which prices. Moreover, analogous to (2), the seller's total profit is a weighted sum of one-period profits from a buyer who knows her value to be a cutoff price. We complete the proof by showing that these cutoff prices are less dispersed (in the sense of SOSD) than the buyer's expected values at the end of T periods. Thus there exists a static information structure that induces these cutoff prices as posterior expected values, and it yields the same profit as the original dynamic information process.[22]

We mention that Π_{RSD} per buyer can no longer be guaranteed when information exhibits limited price-dependence.[23] We derive this result in Appendix B.4 by considering another variant of our main model, where information depends on past (but not current) prices. More generally, note that limited price-dependence is only one kind of restriction on nature's choice of information structures. In the next section we consider several other reasonable restrictions and characterize the corresponding optimal pricing strategies.

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. As mentioned in the introduction, one motivation for such a result is to use it as a benchmark, so that we can explain certain selling strategies as arising optimally from additional qualitative features of the environment. We view this as an advantage of the robust modeling approach, since carrying out the same exercise under a standard Bayesian objective would (at least) require a tractable description of prior beliefs over dynamic information structures.

In this section, we consider several modifications of our main model, where the seller seeks robustness against a *subset* of information arrival processes (with the whole set described in Section 2.1). We find that certain reasonable restrictions on the environment lead to dynamic pricing strategies.[24] In Section 6.1, we show how declining prices out-perform constant prices when

[22]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.

[23]This impossibility result suggests that the seller strictly benefits from personalized pricing.

[24]That said, not all restrictions affect the optimality of constant price paths. As discussed, constant prices remain optimal if the only restriction is that buyers learn according to a partitional information structure.

the seller believes that buyers do not learn in every period (e.g., when information is somewhat rare). In Section 6.2, we consider buyers who arrive with some additional information beyond the knowledge of the prior value distribution. Increasing prices can be optimal if later buyers have better information to begin with. Finally, in Section 6.3 we show that introductory pricing is favored when buyer values are common 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 (for the first buyer) 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: If the product is complicated or marketed on a small scale, buyers may only learn about it from a few particular sources (e.g., in person with a technology expert). In this case, learning may not occur every period.

The following result shows that the seller can now obtain a higher profit guarantee with a decreasing price path. As a corollary, the optimal deterministic pricing strategy involves decreasing prices.

Proposition 3. *Suppose that $T = 2$ and that the first buyer either receives information in period one or period two, but not both. Further suppose $G(p^*) > 0$. Then for any $\delta \in (0, 1)$, there exists a price path $p_1 > p_2 = p^*$ that guarantees profit strictly greater than Π^* from the first buyer.*

Since $p_2 = p^*$ is optimal for the second buyer, the seller's total profit exceeds $(1 + \delta)\Pi^*$.

The intuition for this proposition goes back to the upper-bound argument (Lemma 2) in Section 4.1.2. There we showed how nature can use a partitional information arrival 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.[25] But if nature were to remove the threshold in the second period, then sale would jump up in that period unless $G(p^*) = 0$. 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.

[25] Note that any 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. Initial Information

So far we have assumed that the seller has zero knowledge over what information buyers receive. But in practice, he may know that buyers necessarily observe certain signals. For example, the seller may himself conduct an advertising campaign and understand its informational impact.[26]

This situation can be modeled by assuming that the buyer who arrives at time a observes signal $\underline{s_a}$ according to some *initial information structure* $\mathcal{H}_a$, in addition to knowing the prior F. The seller knows this information structure (i.e., how $\underline{s_a}$ is distributed given v_a), but does not observe the realization of $\underline{s_a}$. We maintain all aspects of the main model, except that we allow nature to provide information conditional on $\underline{s_a}$.[27]

To solve this extension, we first study the static problem with a single buyer. Let $F_{\underline{s_a}}$ be buyer a's posterior value distribution upon observing signal $\underline{s_a}$. The same analysis shows that for this "prior" value distribution, the worst-case static information structure is partitional. Hence, if we let $G_{\underline{s_a}}$ be the pressed distribution of $F_{\underline{s_a}}$, we have the following result:

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

$$p^*_{\mathcal{H}_a} \in \operatorname*{argmax}_p \, p(1 - \mathbb{E}[G_{\underline{s_a}}(p)]), \tag{5}$$

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

In the dynamic model, we obtain a partial characterization of the optimal pricing strategy.

Proposition 4. *Suppose that each buyer a arrives with initial information structure $\mathcal{H}_a$. Further suppose that the one-period optimal prices $p^*_{\mathcal{H}_a}$ defined by (5) increase with a. Then the seller's optimal selling strategy is to charge $p^*_{\mathcal{H}_a}$ in each period a, resulting in a (weakly) increasing price path.*

Note that a special case of the model considered here involves the same initial information structure for all buyers. Proposition 4 implies that a constant price path remains optimal. This generalizes our Theorem 1 as well as Stokey's constant price result, since the known value setting corresponds to perfectly informative initial information.

Another extreme case is as follows: All buyers who arrive prior to some period A have no initial information, while those arriving in period A or afterwards know their values. Here the

[26] If the seller could freely disclose any information, he would provide a perfectly informative signal to minimize the residual uncertainty. A more interesting application would be to study *costly* information disclosure by a seller who is concerned that buyers may receive additional information afterwards. Proposition 4 below would provide a useful first step of the analysis. We mention that Terstiege and Wasser (2018) have studied a related model: Building on Roesler and Szentes (2017), they consider optimal buyer information acquisition that is robust to potentially more information provided by the seller.

[27] Alternatively, we may think of the initial information structure $\mathcal{H}_a$ as a constraint on nature's information choice, meaning that nature's signal in period a must be more informative than $\underline{s_a}$ in the sense of Blackwell (1953).

one-period optimal price for buyer a is either p^* when $a < A$, or the known value monopoly price $\hat{p}$ when $a \geq A$. Since $p^* \leq \hat{p}$ by Lemma 10 in the appendix, Proposition 4 applies and suggests a (discrete) price increase in period A.

6.3. Common Values and Public Information

This section modifies the main model by considering common values and publicly observed signals. Notice that making one change without the other would leave the problem unaltered.[28] Here we argue that with both modifications, the seller is able to guarantee higher profits. 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.

We will study the stylized case of *pure common values* and *perfectly correlated signals*. Specifically, we impose that all buyers have the same value v, which is drawn from F at $t = 0$. Moreover, nature chooses a *single* information arrival process $\mathcal{I}$ that consists of signal sets $(S_t)_{1 \leq t \leq T}$ and signal distributions $I_t : \mathbb{R}_+ \times S^{t-1} \times P^t \to \Delta(S_t)$. As in our main model, signals can depend on past and current prices. However, signals are public, so all buyers in the market have the same information. In particular, a buyer who arrives at time a observes all signals prior to and including period a.[29] Each buyer decides when to purchase based on the signals she observes.[30]

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:*

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

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

Figure 1 below illustrates the price paths we use for this approximation, in the case of a uniform prior. Starting off at Π_{RSD}, prices increase and eventually flatten at a level that converges as $\delta \to 1$ to the number S from (3). Prices also increase more slowly in the patient limit.

[28] To see this, note that a constant price path makes the problem stationary across buyers. So the seller can guarantee the sum of what can be ensured from individual buyers. On the other hand, this total profit guarantee cannot be improved upon with either independent values or private signals, since nature can simultaneously minimize the seller's profit from each buyer.

[29] This setup connects with the previous Section 6.2, since later buyers arrive with more information. The difference is that the amount of initial information is now *endogenous*.

[30] With perfectly-aligned values and signal observations, other buyers' purchase decisions do not reveal more information than the signals. Thus in our setting it is without loss to make purchase decisions unobservable.

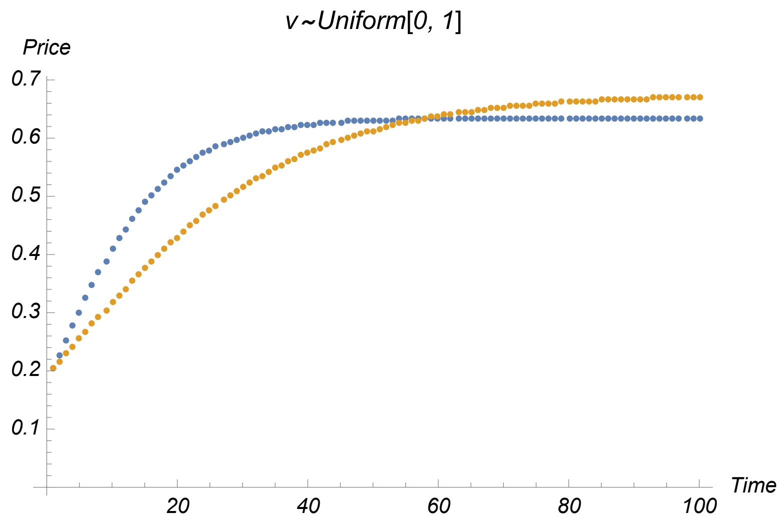

Figure 1: Illustration of price paths. Blue is $\delta = 0.9$; Orange is $\delta = 0.95$.

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 the first period.*

Lemma 4 enables us to restrict attention to static information structures. To complete the proof, we adapt Du's random price distribution (3) to construct price paths such that the seller's total profit under any static information structure approximates the single-period profit under Du's mechanism. As a consequence, profit guarantee converges to Π_{RSD}.

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 to each buyer, and a constant price path delivers this optimal profit. We have shown how this conclusion depends on our formulation of the seller's problem, in particular the assumption regarding whether (the seller believes) prices influence information availability. We also identify several economically meaningful restrictions on the informational environment that would lead to gains from non-constant pricing strategies. This illustrates the power of our approach to explain pricing behavior, since performing a Bayesian analysis with general information structures would typically disallow a parsimonious benchmark result similar to our Theorem 1.

We view one contribution of this paper as introducing a robust objective into a dynamic mechanism design problem. Dynamics complicates 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. We hope that the techniques developed in this paper will help other researchers seeking to 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.*

Let $\underline{v}$ be the minimum value in the support of F. For $\beta \in (\underline{v}, \mathbb{E}[v]]$, define $G(\beta) = \sup\{\alpha : g(\alpha) \leq \beta\}$. We extend the domain of this inverse function to $\mathbb{R}_+$ by setting $G(\beta) = 0$ for $\beta \leq \underline{v}$ and $G(\beta) = 1$ for $\beta > \mathbb{E}[v]$.[31]

The rest of this appendix provide proofs for Proposition 1, Theorem 1 and Proposition 2.

A.1. Proof of Proposition 1

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

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

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

A.2. Proof of Theorem 1

As discussed in the main text, the proof consists of a lower-bound and an upper-bound. We address them in turn and discuss at the end why the optimal pricing strategy is unique (when p^* is).

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

A.2.1. Lower-bound: Proof of Lemma 1

For the lower-bound on the seller's profit guarantee, we will prove Lemma 1 which is stronger. 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 $S = \{\overline{s}, \underline{s}\}$, 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 $\overline{s}$ with probability $\delta^{\tau-1}$. With complementary probability and when $\tau = \infty$, let her receive the other signal $\underline{s}$. In the alternative information structure $\mathcal{I}'$, nature reveals $\overline{s}$ or $\underline{s}$ in the first period and provides no more information afterwards.

We claim that under $\mathcal{I}'$, the buyer receiving the signal $\underline{s}$ has expected value at most p_1. In fact, something stronger holds, namely that the buyer has expected value at most p_1 conditional on $\underline{s}$ and *any* realized s_1 under the original information structure $\mathcal{I}$.[32] To prove this, note that since stopping at time τ is weakly better than stopping at time 1, we have

$$\mathbb{E}[v \mid s_1] - p_1 \leq \mathbb{E}^{s_2,\cdots,s_T}\left[\delta^{\tau-1}(\mathbb{E}[v \mid s_1, s_2, \cdots, s_\tau] - p_\tau)\right]. \tag{6}$$

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

Since $p_\tau \geq p_1$, simple algebra reduces (6) to the following.

$$\mathbb{E}[v \mid s_1] \leq \mathbb{E}^{s_2,\cdots,s_T}\left[\delta^{\tau-1}\mathbb{E}[v \mid s_1, s_2, \cdots, s_\tau] + (1-\delta^{\tau-1})p_1\right]. \tag{7}$$

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

$$p_1 \geq \frac{\mathbb{E}^{s_2,\cdots,s_T}[(1-\delta^{\tau-1}) \cdot \mathbb{E}[v \mid s_1, s_2, \cdots, s_\tau]]}{\mathbb{E}^{s_2,\cdots,s_T}[1-\delta^{\tau-1}]}. \tag{8}$$

The denominator $\mathbb{E}^{s_2,\cdots,s_T}[1-\delta^{\tau-1}]$ can be rewritten as $\mathbb{E}^{s_2,\cdots,s_T}\left[\mathbb{P}(\underline{s} \mid s_1, s_2, \ldots, s_T)\right]$, which

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

is the probability of $\underline{s}$ given s_1. Because τ is a stopping time, the numerator in (8) equals

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

which can be further rewritten as

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

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

With these, (8) states that

$$p_1 \geq \frac{\mathbb{E}^{s_2,\cdots,s_T}\left[\mathbb{P}(\underline{s}\mid s_1,s_2,\ldots,s_T)\cdot\mathbb{E}[v\mid s_1,s_2,\cdots,s_T,\underline{s}]\right]}{\mathbb{E}^{s_2,\cdots,s_T}\left[\mathbb{P}(\underline{s}\mid s_1,s_2,\ldots,s_T)\right]} = \mathbb{E}[v|s_1,\underline{s}] \tag{9}$$

just as we claimed.

Thus, under the static information structure $\mathcal{I}'$ constructed above, a buyer who receives the signal $\underline{s}$ has expected value at most p_1, which is also less than any future price. Since information only arrives in the first period, all sale happens in the first period to the buyer receiving the signal $\overline{s}$. The probability of sale is at most $\mathbb{E}[\delta^{\tau-1}]$, and the seller's profit is at most $\mathbb{E}[\delta^{\tau-1}]\cdot p_1$. This is no more than $\mathbb{E}[\delta^{\tau-1}\cdot p_\tau]$, the discounted profit under the original dynamic information structure. Hence the lemma.

A.2.2. 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

$$v - p_t \geq \max_{\tau\geq t+1}\mathbb{E}[\delta^{\tau-t}\cdot(v-p_\tau)] \tag{10}$$

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 $v\geq v_t$.[33] Thus, v_t is defined by the equation

$$v_t - p_t = \max_{\tau\geq t+1}\mathbb{E}[\delta^{\tau-t}\cdot(v_t-p_\tau)] \tag{11}$$

and it is a random variable that depends on realized prices p^t and the expected distribution of

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

future prices $\sigma(\cdot \mid p^t)$.

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

$$w_t = \min\{v_1, v_2, \ldots, v_t\} = \min\{w_{t-1}, v_t\}. \tag{12}$$

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

Step 3: Buyer behavior. 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 structure $\mathcal{I}$ be constructed as above. Then the first 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 (12). Given this 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. The lemma follows.

Step 4: Profit decomposition. By Lemma 5, the buyer whose true value belongs to the percentile range $(G(w_{t-1}), G(w_t)]$ 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*

$$p_t = \mathbb{E}\left[\sum_{s=t}^{T-1}(1-\delta)\delta^{s-t}w_s + \delta^{T-t}w_T \mid p^t\right] \tag{13}$$

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

Using Lemma 6, we can rewrite the profit as

$$\begin{aligned}
\Pi &= \mathbb{E}\left[\sum_{t=1}^{T}\delta^{t-1}\cdot(G(w_{t-1})-G(w_t))\cdot\mathbb{E}\left[\sum_{s=t}^{T-1}(1-\delta)\delta^{s-t}w_s+\delta^{T-t}w_T \mid p^t\right]\right]\\
&= \mathbb{E}\left[\sum_{t=1}^{T}\delta^{t-1}\cdot(G(w_{t-1})-G(w_t))\cdot\left(\sum_{s=t}^{T-1}(1-\delta)\delta^{s-t}w_s+\delta^{T-t}w_T\right)\right]\\
&= \mathbb{E}\left[\sum_{s=1}^{T-1}(1-\delta)\delta^{s-1}w_s(1-G(w_s))+\delta^{T-1}w_T(1-G(w_T))\right]\\
&\leq \Pi^*.
\end{aligned} \tag{14}$$

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,[34] 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 (11) we can find an optimal stopping time $\tau \geq t+1$ such that

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

which can be rewritten as

$$p_t = \mathbb{E}[(1-\delta^{\tau-t})v_t + \delta^{\tau-t}p_\tau]. \tag{15}$$

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

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

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 (15) proves the lemma as well as Theorem 1.

A.2.3. Proof of Unique Optimality

Suppose p^* is unique in the one-period problem, then from (14) we see that the seller's profit from the first buyer equals Π^* only if $w_s = p^*$ almost surely for each s. This together with Lemma 6 implies $p_1 = p^*$ with probability 1. Similar consideration for later buyers shows that the seller must always charge p^* to achieve the total profit guarantee $\Pi^* \cdot \frac{1-\delta^T}{1-\delta}$.

A.2.4. Example: Profit Can be Even Worse

The partitional information structure 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:

Example 2. *Let $T = 2$, $v = 0$ or 1 with equal probabilities, and $\delta = 1/2$. Suppose the seller sets prices to be $p_1 = 11/40$ and $p_2 = 1/10$. Under these prices, a buyer with value $\frac{9}{20}$ would be indifferent (in the first period) between purchase and delay. Hence the partitional information structure constructed before Lemma 5 induces expected value $\frac{9}{20}$ when recommending the buyer not to purchase in the first period. This information structure 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 structure is

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

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

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

which in fact holds with equality. Under this different information structure, 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. Proof of Proposition 2

Fix any dynamic information arrival process, our goal is to find a *partitional* process that leads to lower profit. For ease of exposition, we present the proof assuming a deterministic price path, but the same argument applies to randomized pricing strategies.

As a first step, we assume without loss that the original process $\mathcal{I}$ simply recommends the buyer to purchase or not in each period. For $1 \leq t \leq T$, let λ_t denote the probability that the buyer is recommended to purchase in period t, and let y_t denote her expected value given this recommendation (and previous recommendations not to purchase). We also define λ_{T+1} and y_{T+1} to correspond to the situation when the buyer is never recommended to purchase.

In the alternative, partitional, process $\mathcal{I}'$, we consider thresholds $\infty = v_0 \geq v_1 \geq \cdots \geq v_T \geq v_{T+1} = \underline{v}$, such that $\mathbb{P}[v \in [v_t, v_{t-1})] = \lambda_t$. Under this partitional process, the buyer learns whether or not $v \geq v_t$ in each period t. We use z_t to denote the average value when v belongs to the interval $[v_t, v_{t-1})$. Crucially, we have the following inequality

$$\sum_{r=t+1}^{T+1} \lambda_r \cdot y_r \geq \sum_{r=t+1}^{T+1} \lambda_r \cdot z_r, \quad \forall 0 \leq t \leq T. \tag{16}$$

This reflects a key property of partititional information structures: given a mass $\sum_{r>t} \lambda_r$ of buyers, their average value is minimized when they are precisely those buyers with value less than v_t.

Using (16), 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

$$\sum_{s=t+1}^{T+1} \lambda_s \cdot (y_s - p_t) \leq \sum_{s=t+1}^{T} \delta^{s-t} \lambda_s \cdot (y_s - p_s).$$

Rearranging, this yields

$$\sum_{s=t+1}^{T} (1 - \delta^{s-t}) \lambda_s y_s + \lambda_{T+1} y_{T+1} \leq \sum_{s=t+1}^{T+1} \lambda_s p_t - \sum_{s=t+1}^{T} \delta^{s-t} \lambda_s p_s.$$

Observe that the LHS above is a positive linear combination of the LHS of (16), so we can use (16) to further deduce (with z_s replacing y_s everywhere)

$$\sum_{s=t+1}^{T} (1 - \delta^{s-t}) \lambda_s z_s + \lambda_{T+1} z_{T+1} \leq \sum_{s=t+1}^{T+1} \lambda_s p_t - \sum_{s=t+1}^{T} \delta^{s-t} \lambda_s p_s.$$

Rearranging again gives

$$\sum_{s=t+1}^{T+1} \lambda_s \cdot (z_s - p_t) \leq \sum_{s=t+1}^{T} \delta^{s-t} \lambda_s \cdot (z_s - p_s).$$

That is, a buyer with value below v_t should not purchase in period t.

By the above analysis, the partitional process $\mathcal{I}'$ ensures that any buyer with value in $[v_t, v_{t-1})$ purchases in period t or later. If she indeed purchases in period t, discounted profit equals $\delta^{t-1} \lambda_t p_t$, which is the same as the original discounted profit from period t. But if she delays, discounted profit would be even lower because social surplus decreases while buyer surplus could only increase. This proves that the constructed partitional process yields a lower profit.

B. Proofs for The Model with Limited Price-dependence

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. We conclude with a discussion of the dynamic model with *limited* price-dependence, in particular focusing on the case where information depends on *past but not current* prices.

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.:

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

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

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

where F_W^B is the Roesler-Szentes worst-case information structure. To explain where S comes from, note that the LHS in (18) must not exceed the RHS *for all* S because F is a mean-preserving spread of F_W^B (Rothschild and Stiglitz (1970)). When W is smallest possible, such a constraint must bind at some S.

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

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

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

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

The second inequality follows because F is a mean-preserving spread of $\tilde{F}$. The next equality uses

(18), and the last equality uses (4).

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 upon arrival, while delivering lower profit to the seller.

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

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

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

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:

$$v_t = \mathbb{E}\left[\sum_{s \geq t} (1-\delta)\delta^{s-t} q_s \mid v_1, \dots, v_t\right]. \tag{19}$$

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

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

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

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

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 (with one buyer) under the dynamic information process is also the profit in one period facing 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 in second-order stochastic dominance. For each real number x, let γ be a stopping time adapted to the v-process such that q_γ first exceeds x. Then

$$\begin{aligned}
\mathbb{E}\left[\sum_{t\geq 1}(1-\delta)\delta^{t-1}(q_t-x)^+\right] &= \mathbb{E}\left[\delta^{\gamma-1}\sum_{t\geq\gamma}(1-\delta)\delta^{t-\gamma}(q_t-x)\right] \\
&= \mathbb{E}\left[\delta^{\gamma-1}(v_\gamma - x)\right] \\
&\leq \mathbb{E}[(v_\infty - x)^+],
\end{aligned}$$

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

B.3. Information Depending on Past Prices

We now consider another version of our model, where information can depend on *past but not current* prices. We show that Π_{RSD} per buyer can still be guaranteed if information is static.

Proposition 6. *Suppose that information to each buyer is only provided when she arrives and can only depend on past prices. Then for any T and δ, maxmin average profit per buyer is Π_{RSD}. The seller can achieve this by a strategy that sets independent prices across periods.*

The proof relies on the following lemma regarding the outcome-equivalence between static and dynamic pricing strategies (under known values):

Lemma 7 (Outcome Equivalence)**.** *Fix any continuous distribution function $D(\cdot)$ and any T, δ.*[35] *There exists a pricing strategy $\sigma \in \Delta(p^T)$ such that in the known value case, any buyer who has true value v and arrives in period a purchases with total probability $D(v)$ (discounted to period a).*

In words, for any *personalized* static pricing strategy D, there is a dynamic pricing strategy σ which does not condition on buyers' arrival times, but which results in the same discounted purchase probabilities for *every type of each arriving buyer.* Note that in Proposition 6, we assume buyers do know their values upon arrival (given by the static information structure). As a consequence, a seller using strategy σ obtains the same profit from each buyer as if he sells only once to this buyer at a random price drawn from D. So Lemma 7 will imply Proposition 6.

Proof of Lemma 7. We first prove the result for $T = 2$, then generalize to all finite T and lastly discuss the case of $T = \infty$.

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

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

$$v_1 - p_1 = \delta \cdot \mathbb{E}^{p_2 \sim D}\left[\max\{v_1 - p_2, 0\}\right]. \tag{20}$$

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

If the buyer has value at least w_1, she purchases in the first period. If here value belongs to $(w_2, w_1]$, she purchases in the second period. Otherwise she does not purchase. The total discounted purchase probability of buyer 1 with value x is thus

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

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

Suppose H is the c.d.f. of v_1. Since $w_1 = v_1$ and $w_2 = \min\{v_1, p_2\}$, the probability that w is *greater than* x is given by $(1-\delta)(1 - H(x)) + \delta(1 - H(x))(1 - D(x))$.[36] This has to be equal to $1 - D(x)$, which implies

$$1 - H(x) = \frac{1 - D(x)}{1 - \delta D(x)}. \tag{21}$$

[35]Note that Du's distribution $D(\cdot)$ is continuous except when it is a point-mass on W; in that exceptional case we have $\Pi_{RSD} = \Pi^*$, and Proposition 6 follows from Theorem 1.

[36]$1 - H(x)$ is the probability that $w_1 > x$, and $(1 - H(x))(1 - D(x))$ is the probability that $w_2 > x$.

We are left with the task of finding a first-period price distribution under which $v_1 \sim H$. This can be done because the random variables v_1 and p_1 are in a one-to-one relation (see (20)). The lemma thus holds for $T = 2$.

Remark 2. *It will be useful to note that (21) implies the distribution H has the same support as D. In light of (20), we see that when v_1 achieves the maximum of this support, p_1 is in general strictly smaller than v_1 (unless the support is a singleton point, which implies $\Pi_{RSD} = \Pi^*$). So the first-period maximum price is smaller than in the second period. On the other hand, the minimum price p_1 is equal to the minimum of the support of D, which is W when D is Du's price distribution.*

Step 2: Extension to finite T. Similar to the above, we conjecture a pricing strategy σ that is independent across periods: $d\sigma(p_1, \dots, p_T) = d\sigma_1(p_1) \times \cdots \times d\sigma_T(p_T)$, where we interpret each σ_t as a distribution. Define the cutoff values $v_1, \dots, v_T$ as in (11). Note that due to independence, v_t only depends on current price p_t but not on previous prices. Further define random variables $w^{(t)}$ as follows: For $t \le s \le T-1$, $w^{(t)} = \min\{v_t, v_{t+1}, \dots, v_s\}$ with probability $(1-\delta)\delta^{s-t}$; and with remaining probability δ^{T-t}, $w^{(t)} = \min\{v_t, v_{t+1}, \dots, v_T\}$.

Consider a buyer who arrives in period t with value x. We can generalize the previous arguments and show that her discounted purchase probability is $\mathbb{P}[w^{(t)} \le x]$. To deduce the lemma, we want each $w^{(t)}$ to be distributed according to D. Simple calculation shows this is the case if $v_T \sim D$ and $v_1, \dots, v_{T-1} \sim H$ as given in (21); since v_t depends only on p_t, they are independent random variables. We can then solve for the price distributions $\sigma_1, \dots, \sigma_T$ by backward induction: σ_T is simply D, and once the prices in period $t+1, \dots, T$ are determined, there is a one-to-one relation between p_t and v_t by (11). Thus, the distribution of p_t is uniquely pinned down by the distribution of v_t (which we know is H).

Step 3: The infinite horizon case. If $T = \infty$, we look for price distributions $\sigma_1, \sigma_2, \dots$ such that $v_1, v_2, \dots \sim H$. The reason $H(x)$ should be the c.d.f. of v_1 can be understood this way: Under stationary price distributions, the (first) buyer with value x purchases in period t with probability $H(x)$ conditional on not purchasing earlier. Thus the discounted purchase probability is $\sum_t \delta^{t-1}(1 - H(x))^{t-1} H(x)$. Setting this equal to $D(x)$ yields (21).

In order for each v_t to be distributed as H, we conjecture a stationary price distribution. Below we let $P(x)$ denote the c.d.f. of each p_t. To determine P, recall that the cutoff v_1 is defined by

$$v_1 - p_1 = \max_{\tau \ge 2} \mathbb{E}\left[\delta^{\tau-1}(v_1 - p_\tau)\right]. \tag{22}$$

The stopping problem on the RHS is stationary. Thus when $p_2 < p_1$ the buyer stops in period 2 and receives $v_1 - p_2$; otherwise she continues and receives $v_1 - p_1$. (22) thus reduces to

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

which can be further simplified to

$$v_1 = p_1 + \frac{\delta}{1-\delta} \cdot \mathbb{E}^{p_2}\left[\max\{p_1 - p_2, 0\}\right]. \tag{23}$$

When $p_1 = x$, (23) implies

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

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

$$P(x) = H\left(x + \frac{\delta}{1-\delta} \int_0^x P(z)\, dz\right). \tag{24}$$

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

$$Q(x) = x + \frac{\delta}{1-\delta} \int_0^x P(z)\, dz; \qquad U(y) = 1 + \frac{\delta}{1-\delta} H(y) = \frac{1}{1 - \delta D(y)}. \tag{25}$$

Then (24) becomes the differential equation

$$U(Q(x)) = Q'(x). \tag{26}$$

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

$$\frac{\partial V(Q(x))}{\partial x} = V'(Q(x)) \cdot Q'(x) = \frac{Q'(x)}{U(Q(x))} = 1. \tag{27}$$

Inspired by the analysis for finite T, we conjecture that the minimum value of p_1 is W. That is, we conjecture $Q(W) = W$. Thus $V(Q(W)) = V(W) = W$. From (27) we deduce $V(Q(x)) = x$, so that

$$Q(x) = V^{-1}(x) \text{ with } V(y) = \int_0^y (1 - \delta D(z))\, dz. \tag{28}$$

Since V is strictly increasing, there is a unique solution $Q(x)$ to the above equation. From (25), the corresponding distribution of prices is

$$P(x) = \frac{1-\delta}{\delta} \cdot (Q'(x) - 1). \tag{29}$$

This completes the proof of Lemma 7 and thus of Proposition 6.

B.4. An Impossibility Result

Finally, consider information structures that are *dynamic and depend on past but not current prices.* Π_{RSD} remains an upper-bound on the seller's profit guarantee. However, unlike in the previous section, we show here that the upper-bound cannot be achieved against dynamic information. The following result is stated for two periods, but it easily generalizes to any $T > 1$.

Proposition 7. *Suppose that information can depend on past prices,* $\Pi_{RSD} > \Pi^*$ *and Du's mechanism is uniquely maxmin optimal in the one-period problem.*[37] *Then in the two-period model with one buyer arriving in each period, the seller's total discounted profit guarantee is strictly below* $(1+\delta)\Pi_{RSD}$ *for any* $\delta \in (0, 1)$.

Proof. We present the proof in several steps.

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

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

- In the first period, nature provides the Roesler-Szentes information structure. We denote the buyer's posterior expected value by $\tilde{v}$, so as to distinguish from her true value v. Note that $\tilde{v} \sim F_W^B$.
- In the second period, given the realized price p_1 and the buyer's expected value $\tilde{v}$ after the first period, nature reveals the buyer's true value v if $\tilde{v} \geq v_1(p_1)$, where the cutoff $v_1(p_1)$ is defined as usual (assuming no information arrives in the second period):

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

  If $\tilde{v} < v_1(p_1)$, nature provides no information in the second period.

Intuitively, nature targets the buyer who prefers to purchase in the first period *when she does not expect to receive information in the second period.* By promising full information to such a buyer in the future, nature potentially delays her purchase and reduces the seller's profit. The rest of the

[37] $\Pi_{RSD} > \Pi^*$ is clearly necessary for the result: We have shown in our main model that $(1+\delta)\Pi^*$ can be guaranteed even when nature is allowed to condition on current price. On the other hand, we impose an extra assumption that Du's mechanism is strictly optimal. This is for technical reasons that we explain below, and it may not be necessary for the conclusion. In any event, we show in Appendix D.4 that Du's mechanism is indeed unique for "generic" value distributions F.

proof formalizes this argument.

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

Lemma 8 (Profit Comparison)**.** *Consider the dynamic information structure $\mathcal{I}$ and its static counterpart $\mathcal{I}'$ constructed above. The seller's profit under $\mathcal{I}'$ is no greater than Π_{RSD}, and his profit under $\mathcal{I}$ is even smaller by at least $(1-\delta)W$ multiplied by the probability that the buyer delays purchase.*

To prove this lemma, we consider three possibilities. First, if the price p_1 is so high that $\tilde{v} < v_1(p_1)$, then the buyer does not purchase in the first period under $\mathcal{I}'$. This is also her optimal decision under $\mathcal{I}$, because she will not receive extra information in the second period. Second, if the price is very low, then under both $\mathcal{I}$ and $\mathcal{I}'$ the buyer purchases in the first period. Lastly, *for some intermediate prices*, the buyer purchases in the first period under $\mathcal{I}'$ but not under $\mathcal{I}$. We note that the opposite case cannot arise, because $\mathcal{I}$ provides more information than $\mathcal{I}'$ in the second period, making the buyer more willing to delay under $\mathcal{I}$.

To summarize, when nature provides the dynamic information structure $\mathcal{I}$ rather than $\mathcal{I}'$, the seller's profit changes only when some buyers delay their purchase in the first period. Observe that whenever such delay occurs, *discounted social surplus* decreases from $\tilde{v}$ to at most $\delta \cdot \tilde{v}$. Since buyer surplus cannot decrease by incentive compatibility, profit must decrease by at least $(1-\delta)\tilde{v} \geq (1-\delta)W$. Lemma 8 follows.

Step 3: Proof for a particular pricing strategy. Let σ^D be the pricing strategy given by Lemma 7, which we recall guarantees Π_{RSD} from each buyer when information is static. Here we argue that if the seller uses σ^D, then by providing the dynamic information structure $\mathcal{I}$, nature holds profit strictly less than Π_{RSD}. Later we generalize the result to other pricing strategies.

Remark 2 shows that under σ^D, p_2 is drawn from Du's distribution D independently of p_1. On the other hand, p_1 is continuously supported on a smaller interval $[W, S_1]$, with $W < S_1 < S$. The distribution of p_1 is such that $v_1(p_1) \sim H$, which is supported on $[W, S]$.

Suppose the buyer's posterior expected value $\tilde{v}$ belongs to the open interval (W, S). Further suppose that knowing her true value *strictly* improves her expected payoff in the second period when $p_2 \sim D$. Then, whenever p_1 is smaller than but close to the indifference type $v_1^{-1}(\tilde{v})$, such a buyer would purchase in the first period under $\mathcal{I}'$ but delay purchase under $\mathcal{I}$. By Lemma 8, we can bound profit away from Π_{RSD} so long as we find a positive measure of such buyers.

To do this, note from Remark 1 that $\tilde{v} < S$ implies the true value also satisfies $v < S$. Moreover, because we assume $\Pi_{RSD} > \Pi^*$, Lemma 10 in Appendix D.5 gives $W > \underline{v}$. Thus with positive probability, a buyer with expected value $\tilde{v} \in (W, S)$ has true value $v \in (\underline{v}, W)$. For any such buyer, even without additional information she would purchase at some prices $p_2 \sim D$ and $p_2 \in (W, \tilde{v})$. But if she were informed that her true value is less than W, she would not purchase at any second-period price $p_2 \sim D$. Hence knowing her true value strictly improves this buyer's second-period expected payoff, and we are done with the proof when the seller uses σ^D.

Step 4: Proof for general pricing strategy σ. We now prove Proposition 7 in its full generality. The argument is as follows (omitting some technical details): Suppose for contradiction that some pricing strategy σ guarantees profit arbitrarily close to Π_{RSD} from each buyer. Then because Du's price distribution $D(x)$ is *uniquely optimal* in the one-period problem, the distribution of p_2 conditional on p_1 is close to D (in the Prokhorov metric) with high probability; otherwise nature could sufficiently damage the seller's profit from the second buyer. Next, we can similarly show that the distribution of $v_1(p_1)$ must be close to H, which is its distribution under σ^D.[38] The rest of the proof proceeds as in Step 3: With positive probability the buyer has true value $v < W$ and posterior expected value $\tilde{v} \in (W, S)$. For such a buyer, full information in the second period is strictly valuable, and she delays purchase with positive probability under the dynamic information structure $\mathcal{I}$ (relative to $\mathcal{I}'$). Lemma 8 then implies that profit from the first buyer is bounded away from Π_{RSD}. This contradiction proves Proposition 7.

C. Proofs for Other Extensions

C.1. Proof of Proposition 3

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

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

given the static information structure. Then profit can be computed as

$$\begin{aligned}\Pi &= p_1(1-\tilde{F}(v_1)) + \delta p_2(\tilde{F}(v_1) - \tilde{F}(v_2)) \\ &= (1-\delta)v_1(1-\tilde{F}(v_1)) + \delta v_2(1-\tilde{F}(v_2)),\end{aligned} \tag{30}$$

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:

$$\int_0^x F(s)\,ds \geq \int_0^x \tilde{F}(s)\,ds \geq (v_1 - p^*)\tilde{F}(p^*) + (x - v_1)\tilde{F}(v_1), \tag{31}$$

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 (31), we deduce the following inequality:

$$\tilde{F}(v_1) - G(p^*) \leq \frac{v_1 - p^*}{x - v_1} \cdot (G(p^*) - \tilde{F}(p^*)). \tag{32}$$

Plugging into the objective function (30), we conclude that for v_1 sufficiently close to p^* and $\epsilon > 0$ sufficiently small:

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

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

$\limsup \tilde{F}(v_1) \le G(p^*) < 1$ from (32). 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 $\frac{\delta}{1-\delta} \ge \frac{v_1 - p^*}{x - v_1}$ and $G(p^*) - \tilde{F}(p^*) \ge 0$, the latter of which follows from (32) and $\tilde{F}(v_1) \ge \tilde{F}(p^*)$.

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

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

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 partitional 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 $\overline{s}$ in the second period that recommends the buyer to purchase. Then we can rewrite the incentive compatibility condition as

$$\mathbb{E}[v] - p_1 \le \delta \cdot \mathbb{P}[\overline{s}] \cdot (\mathbb{E}[v \mid \overline{s}] - p_2). \tag{33}$$

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

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

Therefore we have the following profit lower-bound:

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

where the first inequality uses the IC constraint (33).

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 4

On one hand, the Replacement Lemma implies that when using the increasing price path specified in the proposition, the seller obtain from each buyer what he can guarantee in the static problem. On the other hand, the proof of Lemma 2 extends to show that a longer selling horizon does not improve the seller's profit from any buyer. Combining both parts yields the proposition.

C.3. 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:

$$(1-\delta) \cdot \Pi^C(\delta, T) = \min_{\tilde{F}} \sum_{t=1}^{T} (1-\delta)\delta^{t-1} p_t \cdot (1 - \tilde{F}(p_t)), \tag{34}$$

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 $p_1, \dots, p_T$ 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 $p_1, \dots, p_T$. 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 $p_1, \dots, p_T$ 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. Fixing any (public) dynamic information structure $\mathcal{I}$, we will replace it with another information structure $\mathcal{I}'$ that only provides a single public signal in the first period. Moreover, we will ensure that each arriving buyer has lower discounted purchase probability under this replacement, so that profit is decreases.

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

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

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

Now we are ready to construct the replacement information structure $\mathcal{I}'$. We assume the signal set is $\{0, 1, \ldots, T\}$, where the signal "i" represents nature's recommendation that buyers with $a \leq i$ purchase upon arrival and that other buyers do not purchase. Furthermore, given the original signal history $s_1, s_2, \ldots$, we assume that signal i realizes only if $i = j_m - 1$ for some critical buyer j_m. We specify the probability of such a signal to be[39]

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

To interpret, these probabilities ensure that conditional on $i \geq j_m - 1$, the event $i \geq j_m$ occurs with probability $\delta^{\tau_{j_m} - j_m}$. In other words, the replacement information structure recommends the critical buyer j_m to purchase with *conditional* probability $\delta^{\tau_{j_m} - j_m}$. This is in line with the proof of Lemma 1 since we push and discount nature's recommendation to this buyer's arrival time. Due to conditioning, however, a difference arises here in that $\delta^{\tau_{j_m} - j_m}$ is *not* the probability of receiving a signal $i \geq j_m$ (except for $m = 1$). From the above formula, we see that any critical buyer is recommended to purchase with probability *smaller than* $\delta^{\tau_{j_m} - j_m}$. In fact, this holds also for non-critical buyers.[40] Thus discounted purchase probabilities are lower as long as buyers are willing to follow nature's recommendation *not to purchase* the object.

Suppose buyer a receives signal $i^* < a$, we need to verify that her expected value is lower than p_a. Since all buyers have the same expectation and prices are increasing over time, it is sufficient to consider $a = i^* + 1$. Then by definition, a must be a critical buyer j_m. We will prove a stronger result, that conditional on *any* realizations $s_1, \ldots, s_{j_m}$ (and on the signal i^*), expected value is at most p_a. Indeed, once $s_1, \ldots, s_{j_m}$ are fixed, so are the critical buyers before j_m as well as their stopping times. Thus the term $\delta^{\sum_{k<m} \tau_{j_k} - j_k}$ is simply a multiplicative constant in the probability of those signals $i \geq i^*$. This suggests that the conditional probability of receiving signal i^* is $1 - \delta^{\tau_{j_m} - j_m}$. But then we return to the proof of Lemma 1, where the buyer is recommended to not purchase with probability $1 - \delta^{\tau_{j_m} - j_m}$. From that proof we know that the buyer's expected value upon seeing i^* is at most p_a. Hence the same is true here, completing the proof.

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

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

D. Other Results

D.1. Uncertainty Leads to Lower Price

We prove here that uncertainty over the information structure leads the seller to choose a lower price than under known values.

Proposition 8. *For any continuous distribution F, let $\hat{p}$ be an optimal monopoly price under known values:*

$$\hat{p} \in \operatorname*{argmax}_{p} p(1 - F(p)). \tag{35}$$

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

Proof of Proposition 8. 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,

$$pG(p) = \int_0^{F^{-1}(G(p))} v dF(v), \forall p \in [\underline{v}, \mathbb{E}[v]]. \tag{36}$$

Differentiating both sides with respect to p, we obtain

$$G(p) + pG'(p) = \frac{\partial}{\partial p}(F^{-1}(G(p))) \cdot F^{-1}(G(p)) \cdot F'(F^{-1}(G(p))) = G'(p) \cdot F^{-1}(G(p)). \tag{37}$$

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

$$G'(p) = \frac{G(p)}{F^{-1}(G(p)) - p}. \tag{38}$$

Thus,

$$G(p) + pG'(p) = \frac{G(p) \cdot F^{-1}(G(p))}{F^{-1}(G(p)) - p}. \tag{39}$$

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

$$\frac{p}{1 - G(p)} \cdot \left(1 - F\left(\frac{p}{1 - G(p)}\right)\right) < p. \tag{40}$$

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

D.2. Known Information Arrival Process

This appendix supplies two simple examples where the information arrival process is known to the seller. We show that optimal prices in such a problem can be increasing or decreasing over time. These examples illustrate why, under the standard Bayesian approach, it may be difficult to accommodate learning while obtaining clear predictions on pricing.

Example 3. *In the first example, consider $F = U[0,1]$, $T = 2$ and $\delta = 1/2$. In the first period no information is provided to the first buyer, while* both *buyers learn their exact value in the second period. Given a price $p_2 \leq 1$, the first buyer purchases upon arrival if $\frac{1}{2} - p_1 \geq \int_{p_2}^{1}(v - p_2)dv = \frac{(1-p_2)^2}{2}$. So if the seller desires to have sale in the first period, it is best to set this indifference condition to hold with equality. In this case profit is $p_1 + \delta p_2(1-p_2) = \frac{1}{2} - \frac{(1-p_2)^2}{2} + \frac{p_2(1-p_2)}{2}$. Optimizing over p_2 gives $p_2 = \frac{3}{4}$, $p_1 = \frac{15}{32}$, yielding profit $\frac{9}{16}$.*

Note that this profit is higher than selling only in the first period (profit $1/2$) or selling only in the second period (profit $1/4$). Hence the optimal (deterministic) price path is $\frac{15}{32}$ followed by $\frac{3}{4}$, which is increasing over time. We mention that the seller does not benefit from randomizing in the second period. In fact, when p_2 is a random variable, p_1 is bounded above by $\frac{1}{2} - \mathbb{E}[\frac{(1-p_2)^2}{2}]$. So profit is at most $\frac{1}{2} - \mathbb{E}[\frac{(1-p_2)^2}{2}] + \mathbb{E}[\frac{p_2(1-p_2)}{2}]$. This is maximized by choosing $p_2 = \frac{3}{4}$ with probability one.

Example 4. *In this different example, each buyer is one of two types, L or H, with equal probabilities. The type H buyer has value equal to 1. In contrast, the type L buyer has value $\frac{2}{3}$ with probability $3/4$ and value 0 with probability $1/4$. The information structure is such that in the first period, the first buyer only knows her type, while both buyers know their exact values in the second period.*

Using similar arguments to the previous example, we can show that the (uniquely) optimal prices are $p_1 = 1 - \frac{\delta}{3}$ and $p_2 = \frac{2}{3}$. This choice ensures that when the first buyer has type H, he is indifferent between purchasing in either period. Moreover, $p_2 = \frac{2}{3}$ generates maximal profit when the first buyer has type L, and this price is also optimal facing the second buyer. Hence the optimal price path involves $p_1 > p_2$ and the seller benefits from intertemporal price discrimination.

The intuitive difference between these two examples is that in the first, the seller has the possibility of extracting full surplus from the first buyer. This gives him the incentive to set a higher price in the second period, so that the first buyer does not delay purchase despite future information. In the second example, however, the seller's monopoly profit (a la Stokey) actually increases from $\frac{1}{2}$ to $\frac{7}{12}$ when a buyer learns her true value. This motivates the seller to screen different types of buyers by using a declining price path. Under general learning processes, either or both of these forces may be at play, which explains why the Bayesian approach typically do not provide a sharp prediction.

D.3. 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 F_W^B. 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 $\mathcal{I} : V \times P \to \Delta(S)$.[41]

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 Nash equilibrium of this game, the seller's profit is* Π^* *and the buyer's expected payoff is* $\mathbb{E}[v] - \Pi^*$.

Similar to Roesler and Szentes (2017), trade occurs with probability 1 in equilibrium. However, since in our main model trade is inefficient, the buyer's payoff is higher in this game than in the worst-case scenario for the seller.

Proof of Proposition 9. For each price p, let $\mathcal{I}^*(p)$ be the corresponding worst-case partitional information structure in our main model. We first construct a (subgame-perfect) equilibrium as follows: On the equilibrium path, the buyer chooses to acquire no information if $p = \Pi^*$, but for any other price he acquires information according to $\mathcal{I}^*(p)$; the seller chooses $p = \Pi^*$. Off the equilibrium path, the buyer chooses any different information structure and the seller best responds with some price.

To see this is an equilibrium, observe that on path, trade occurs with probability 1 because $\Pi^* < \mathbb{E}[v]$ whenever F is non-degenerate. Hence the seller's profit is Π^* and the buyer's payoff is

[41] 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).

$\mathbb{E}[v] - \Pi^*$, sharing all the surplus. By the definition of Π^*, choosing $p = \Pi^*$ is the seller's best response on the equilibrium path. It remains to check that the buyer cannot profitably deviate. Indeed, regardless of the buyer's choice of information structure, the seller can always set price to be p^* and guarantee profit Π^*. Since the seller best responds, his actual profit must be higher. But total surplus cannot exceed $\mathbb{E}[v]$, which implies that buyer's payoff is at most $\mathbb{E}[v] - \Pi^*$. This verifies our equilibrium construction.

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

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

D.4. Uniqueness of Du's Mechanism

Recall the random price mechanism from Section 5 and further discussed in Appendix B.1. In general, there could be more than one point S for which (18) 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.[42] The intuition is simple: (18) must bind at *some* S when W is smallest possible (subject to F being a mean-preserving spread of F_W^B). But for (18) 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 9. *There is a uniquely-optimal random price distribution in the one-period price-independent model if and only if (18) holds at a unique point S.*

Proof of Lemma 9. "Only if" follows from 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

[42] 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 $x < x_0$, (18) implies S cannot be the smaller root x_0. Hence S must be the bigger root x_1.

according to $D(\cdot)$. Let $h(p)$ be the p.d.f. of the random price, then seller's profit is given by

$$\Pi = \int_0^1 p \cdot h(p) \cdot (1 - \tilde{F}(p))\, dp. \tag{41}$$

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 $p \cdot h(p) > p' \cdot h(p')$ 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 9 follows.

D.5. 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 10. $\Pi_{RSD} \geq \Pi^*$ *with equality if and only if* $W = \underline{v} = p^*$. *Furthermore, as the distribution* F *varies, the ratio* Π_{RSD}/Π^* *is unbounded.*

Proof of Lemma 10. 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.

References

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

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

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

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

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.

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

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

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

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

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

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

Yiwei Chen and Vivek Farias. Robust dynamic pricing with strategic customers. *Mathematics of Operations Research*, Forthcoming, 2018.

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

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

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

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.

Francesc Dilmé and Daniel Garrett. A dynamic theory of random price discounts. Working Paper, University of Bonn and Toulouse School of Economics, 2017.

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*, Forthcoming, 2018.

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.

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

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

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

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

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.

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.

Anne-Katrin Roesler and Balasz 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.

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

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

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