

Informational Robustness in Intertemporal Pricing

JONATHAN LIBGOBER AND XIAOSHENG MU

Department of Economics, Harvard University

October 26, 2017

ABSTRACT. Consumers may be unsure of their willingness-to-pay for a product if they are unfamiliar with some of its features or have never made a similar purchase before. How does this possibility influence optimal pricing? To answer this question, we introduce a dynamic pricing model where buyers have the ability to learn about their value for a product over time. A seller commits to a pricing strategy, while buyers arrive exogenously and decide when to make a one-time purchase. The seller does not know how each buyer learns about his value for the product, and seeks to maximize profits against the worst-case information arrival processes. With only a single quality level and no known informational externalities, a constant price path delivers the optimal profit, which is also the optimal profit in an environment where buyers cannot delay. We then demonstrate that introductory pricing can be beneficial when the seller knows information is conveyed across buyers, and that intertemporal incentives arise when there are gradations in quality.

Suppose a monopolist has invented a new durable product, and is deciding how to set prices over time to maximize profit. Consulting the literature on intertemporal pricing,¹ the monopolist would find that keeping the price fixed (at the single-period profit maximizing price) is an optimal strategy when consumers understand the product perfectly (as long as willingness-to-pay is not

CONTACT. jlibgober@fas.harvard.edu, xiaoshengmu@fas.harvard.edu. We are particularly indebted to Drew Fudenberg for guidance and encouragement. We also thank Vivek Bhattacharya, Gabriel Carroll, Mira Frick, Ed Glaeser, Claudia Goldin, Ben Golub, Jerry Green, Oliver Hart, Richard Holden, Johannes Hörner, Ryota Iijima, Yuhta Ishii, Michihiro Kandori, Scott Kominers, Eric Maskin, Doron Ravid, Ben Roth, Hugo Sonnenschein, Tomasz Strzalecki, Elie Tamer, Juuso Toikka, and our seminar audiences at Harvard and MIT for comments. Errors are ours.

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

too time dependent). But a wrinkle arises if consumers may learn something influencing how much they like the product after pricing decisions have been made, a salient issue since the monopolist's product is completely new. For example, when the Apple Watch, Amazon Echo, and Google Glass were released, most consumers had little prior experience to inform their willingness-to-pay. In such a situation, the monopolist might suspect that the buyers' purchasing decisions will depend on the available information—e.g., journalist reviews about the product—which may in turn depend on the chosen prices. The potential for information arrival presents a challenge to the monopolist's pricing problem.

In isolation, components of this setting have been studied extensively. The literature on advertising, for instance, has considered the value of information for new products, treating it as given that there is some information that would inform consumers of their willingness-to-pay (see Bagwell (2007) for a thorough discussion of informative advertising). In the intertemporal pricing literature, Stokey (1979) recognized that willingness-to-pay may change over time, and that such changes can influence the optimal pricing strategy. And other papers on intertemporal pricing, such as Biehl (2001) and Deb (2011), have used exogenous learning by consumers as an informal justification for stochastic changes in value.

Despite this apparent interest, we are not aware of any papers that study dynamic pricing while modeling information arrival explicitly. We suspect one major reason for this absence relates to technical difficulties. While some information arrival processes are straightforward to describe, to do so in general appears to require imposing significant restrictions on the environment. And buyers' decisions depend on the value of information, something that is complicated in static environments and (as far as we are aware) intractable in most general dynamic environments. A more tractable problem could involve dropping the Bayesian updating constraints on the stochastic evolution of buyers' values. While this approach is suitable for studying settings with taste shocks, as done in Deb (2011) or Garrett (2016), it does not fully capture learning. So the question of how to price optimally in the face of information arrival is left unanswered.

In this paper, we wish to explain how various features of a seller's environment relate to the optimality of certain selling strategies. Before doing this, however, it is useful to observe that, empirically speaking, the level of familiarity consumers have with a particular product does not *by itself* appear to dramatically influence a firm's preferred selling strategy. In practice, firms tend to eschew sophisticated pricing strategies, even when consumer learning is significant. Apple regularly markets products using consistent pricing strategies, irrespective of exactly how different or unusual each product is. Amazon utilizes constant prices (sometimes with occasional sales) for new products, such as the smart speaker, as well as products for which there have been many iterations, such as the Fire Tablet. Given this observation, in our view the most useful baseline result for these settings is one where pricing strategies are simple, relative to the

potential complexity of the learning environment. With such a result in hand, we would then feel more confident in using the model to explain the prevalence of certain selling strategies in these environments.

We introduce a model of intertemporal pricing that incorporates dynamic information arrival, and proceed to demonstrate a benchmark result on the optimality of constant price paths. To do this, we adopt the approach of the (quite active) literature on robust mechanism design. A seller commits to a pricing strategy, while buyers observe signals of their values, possibly over time, each according to some information structure (or more precisely, information arrival process). We assume that the seller does not know any part of the information arrival processes that inform the buyers of their values,² and that he commits to a pricing strategy as if the information structures were the worst possible given the pricing decisions. One justification for this worst-case analysis is that the seller may want to guarantee a good outcome, no matter what the information structures actually are.³ For our application, another justification would be that an adversary (e.g. a competitor or disgruntled journalists) may be interested in minimizing the seller's profit. If the firm did not have total control over what information consumers might have access to, our framework would be appropriate.⁴

Prior work utilizing the robust approach has demonstrated how simple mechanisms can be optimal in complex environments when the designer is sufficiently worried about the feature that causes the complexity. As we discuss in the literature review, concern over these features tends to favor mechanisms that are invariant to them, and hence simple. The robust approach is therefore a logical place to look in hopes of obtaining sensible pricing policies as optimal in the presence of information arrival.

As for the commitment assumption, introducing it turns out to be the most straightforward way to analyze the dynamic setting while utilizing the robust mechanism design approach. This assumption, as well as other technical issues which arise due to the combination of dynamics and the maxmin objective, are discussed at length in Section 2.1. For now, we simply comment that firms like Amazon and Apple are widely followed by consumers and industry experts, meaning

²While we assume each buyer knows her entire information arrival process, we show in Appendix D.2 that the results continue to hold if buyers face uncertainty over what information they will receive in the future (with maxmin preferences). This observation highlights that our important assumption is only that buyers are less uncertain than the seller about the information they receive at any time, not that they are less uncertain about information they will receive in the future.

³A more complete discussion of this justification can be found in the robust mechanism design literature, in particular: Chung and Ely (2007), Frankel (2014), Yamashita (2015), Bergemann, Brooks and Morris (2017), Carroll (2015, 2017).

⁴We show in Appendix D.1 that it is not important that the information disclosure policy is set to hurt the seller, as long as it is set to help the buyer. However, the solution in that model is somewhat different than what we describe, and involves consumers always purchasing and no information being released in equilibrium, properties that we find unappealing for our setting. Developing this model to accommodate more realistic features would take us too far afield. Still, this interpretation is inspired by (though distinct from) Roesler and Szentes (2017), which we discuss in depth.

that they are typically able to credibly announce (and stick to) consistent pricing strategies. Commitment is also helpful in allowing us to circumvent issues arising in the spirit of the Coase conjecture, which implies that it is (approximately) worst case for buyers to know their values perfectly.

Our first result is that a longer time horizon does not increase the amount of profit the seller can obtain from each buyer. One explanation is as follows: in each period, the adversary could release information that minimizes the profit in that period. Doing so would make the seller's problem separable across time, eliminating potential gains from decreasing prices. This intuition is incomplete, because the worst-case information structures for different periods need not be consistent, in the sense that past information may prevent the adversary from minimizing profits in the future. While this feature makes it difficult to find the exact worst case for arbitrary price paths, we use partitioned information arrival processes to demonstrate how the adversary can hold the seller to a profit no larger than the single period benchmark.

Although selling only once achieves the optimal profit with a single buyer, this pricing strategy forgoes potential future profit when multiple buyers with i.i.d. values arrive over time. In the classic setting without endogenous information, a constant price path maximizes the profit obtained from each arriving buyer, who either buys immediately upon arrival or not at all. This argument does not extend to our problem, since nature⁵ can induce delay by promising to reveal information to the buyer in the future. Such delay could be costly for the seller, due to discounting. However, we show that as nature attempts to convince the buyer to delay her purchase, it must also promise a greater probability of purchase to satisfy the buyer's incentives. It turns out that, from the seller's perspective, the cost of delayed sale is always offset by the increased probability of sale. We thus show that a constant price path ensures the greatest worst-case profit, and it is in fact strictly optimal with arriving buyers.

We proceed to extend the analysis in two directions. The first direction scrutinizes the technical assumptions which were necessary in order to derive the optimality of stationary pricing. It turns out that a key assumption underlying this sharp result is that information is price-contingent in our baseline model. Any model must take a stand on when learning occurs relative to when prices are set (and realized). When both pricing and learning happen over time, several timing assumptions appear reasonable, and it is not immediately clear to us which assumption best captures reality. We view our main model as the most natural one to study, in part because it is the more cautious benchmark and in part because it more seamlessly extends to dynamic environments. However, we consider a number of alternative setups in Section 7, where we

⁵In the main text, we think of nature choosing the information arrival process to hurt the seller. We find this terminology helpful, but do not necessarily view this literally; we simply use it as a thought experiment to help describe a scenario which results in this particular objective.

show how intertemporal incentives can (or cannot) arise if there is limited interaction between information and prices. We hope that these different modeling choices and corresponding results will be helpful to future researchers in the analysis of other dynamic allocation problems, where information revelation or robustness concerns are significant.

It turns out that a key feature of the environment is that information is price-contingent in our baseline model. When moving to dynamics, a model must take a stand on when learning occurs relative to when prices are set. When both pricing and learning happen over time, it is not immediately clear to us which timing assumption best captures reality. We view our approach as the most natural first step, in part because it is the more cautious benchmark and in part because it more seamlessly extends to dynamic environments. We illustrate this second point by showing how intertemporal incentives can arise if information is limited in how it can depend on the price. We hope that these insights are helpful to other researchers who are attempting to analyze other dynamic allocation problems where information revelation or robustness concerns are significant.

The second direction identifies features of the environment that may force a firm to depart from a simple repetition of the single period optimum. We illustrate this in two main extensions, though as the intertemporal pricing literature is vast, undoubtedly there is more work to be done. First, we add a quality dimension to the seller's product, meaning that consumers must decide both when and what to buy. We show that intertemporal incentives arise in this setting—if the seller utilizes the same menu in every period, then nature can induce delay and give the seller a per-period profit lower than the single period benchmark. Second, when there are informational externalities across buyers, introductory pricing can yield higher profits. This result mirrors others that have been provided in Bayesian settings in the new product pricing literature (e.g., Bose et al. 2006), though our setting differs in that we consider the implications of allowing consumers to delay purchase (with non-myopic buyers).

We begin by reviewing the literature, and then proceed to present the main model. The one period benchmark is studied in Section 3, and we show that intertemporal incentives do not help the seller in Section 4. Using this result, we demonstrate that constant price paths are optimal in Section 5. The remaining sections discuss extensions, demonstrating which features drive the results; in particular, we allow a quality component, an alternative timing assumption and informational externalities across buyers. Section 10 concludes.

1. LITERATURE REVIEW

This paper is part of a large literature that studies pricing under robustness concerns, where the designer may be unsure of some parameter of the buyer's problem. Informational robustness is a special case, and one that has been studied in static settings. The most similar to our one-period

model are Roesler and Szentes (2017) and Du (2017). Both papers consider a setting like ours, where the buyer's value comes from some commonly known distribution, but where the seller does not know the information structure that informs the buyer of her value.⁶ Taken together, these papers characterize the seller's maxmin pricing policy and nature's minmax information structure in the *static* zero-sum game between them.⁷ The one-period version of our model differs from these papers, since we assume that nature can reveal information depending on the *realized* price the buyer faces (see Section 2.1 for further discussion). Moreover, our paper is primarily concerned with dynamics, which is absent from Roesler and Szentes (2017) and Du (2017).

Other papers have considered the case where the value distribution itself is unknown to the seller. For instance, Carrasco et. al. (2017) consider a seller who does not know the distribution of the buyer's value, but who may know some of its moments. If the distribution has two-point support, our one-period model becomes a special case of Carrasco et. al. (2017) in which the seller knows the support as well as the expected value.⁸ 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. This point is elaborated on in Section 9.2.

In our model, nature being able to condition on realized prices is sufficient to eliminate any gains to randomization (even if the randomization is to be done in the future). This may be reminiscent of Bergemann and Schlag (2011), who show (in a one-period model) that a deterministic price is optimal when the seller only knows the true value distribution to be in some neighborhood of distributions.⁹ However, the reasoning in Bergemann and Schlag (2011) is that a single choice by nature yields worst-case profit for all prices. This is not true in our setting, but we are able to construct an information structure for every pricing strategy that shows randomization does not have benefits.

While most of this literature is static, some papers have studied dynamic pricing where the seller does not know the value distribution (as opposed to buyer information structures, as we assume). Handel and Misra (2014) allow for multiple purchases, while Caldentey, Liu, Lobel (2016), Liu (2016) and Chen and Farias (2016) consider the case of durable goods. In our setting,

⁶Du (2017) extends the analysis to a one-period, many-buyer common value auction environment. He constructs a class of mechanisms that extracts full surplus when the number of buyers grows to infinity, despite the presence of informational uncertainty. However, which mechanism achieves the maxmin profit remains an open question for finitely many buyers. This is solved in the special case of two buyers and two value types by Bergemann, Brooks and Morris (2016).

⁷Roesler and Szentes (2017) actually motivate their model as one where the buyer chooses the information structure; they show that this solution also minimizes the seller's profit. See Appendix D.1 for a related interpretation of our model.

⁸These authors do not allow nature to condition on the realized price, so their paper focuses on the alternative timing that we discuss in Section 7.

⁹Their result applies to maxmin profit as in our model. The authors also show that if the seller's objective is instead to minimize regret, then random prices do better.

information arrival places restrictions on *how* the value evolves, and rules out the cases considered in the literature. In addition, these papers look at different seller objectives; the first three study regret minimization, whereas the last one looks at a particular mechanism that approximates the optimum. We highlight that the difference in objective is significant, and avoids a degenerate solution that would arise without additional restrictions on the set of possible value distributions.

The literature on robust mechanism design has become popular in recent years in part due to its ability to provide foundations for the optimality of simple mechanisms, which tend to be observed in practice. For instance, Carroll (2017) shows how uncertainty over the correlation between a buyer's demand for different goods leads to the seller pricing the goods independently.¹⁰ In Carroll (2015), uncertainty over the mapping from an agent's actions into outputs leads to the principal aligning the agent's compensation directly with output. In Frankel (2014), similar alignment arises when there is uncertainty over the agent's bias, and Yamashita (2015) shows how uncertainty about bidders higher order beliefs may favor second price auctions even with interdependent values. At the moment, however, this literature has had much less to say about dynamic environments. Important exceptions are Penta (2015), who considers the implementation of social choice functions in dynamic settings, and Chassang (2013), who shows how dynamics allow a principal to approximate robust contracts which may be infeasible in the presence of liability constraints. As these are both rather different from our setting, we suspect there is much work left to be done in this area.

Several intertemporal pricing papers allow for the value to change over time without explicitly modeling information arrival (absent robustness concerns). Stokey (1979) assumes the value changes deterministically given the initial type. Deb (2014) and Garrett (2016) allow for stochastically changing values, but in these papers the evolution of values violates the martingale condition for expectations.¹¹ The maxmin objective leads us to the study of simple and intuitive information structures, making the buyer's problem tractable. While we believe that a Bayesian version of our problem is worth studying, we are not aware of how to determine a buyer's optimal purchasing behavior under an arbitrary information arrival process.

Finally, it is well known that the literature on informational robustness is related to the literature on information design, which has also recently begun to study dynamics (see Ely (2017) and Ely, Frankel and Kamenica (2015)). While we are ultimately concerned with pricing strategies, this connection is relevant because we describe how a receiver's (buyer) behavior varies depending

¹⁰The general link between dynamic allocations and multi-dimensional screening has been long noted in Bayesian settings (see, for instance, Pavan Segal and Toikka (2014) for a discussion of this point). While it is interesting that we obtain a result that is similar to his, we note that our focus on information arrival and a single-object purchase are significant differences. A more complete formal connection is left to future work.

¹¹Deb (2014) assumes the value is independently redrawn upon Poisson shocks. For Garrett (2016), the value follows a two-type Markov-switching process.

on how a sender (nature) chooses the information structure. This connection allows us to import useful results that have been utilized elsewhere, for instance by Kolotilin, Li, Mylovanov and Zapechelnyuk (2017). In turn, several of our results (in particular, the proof of Lemma 2) bear resemblance to this literature, and they may be of interest outside of our setting.

2. MODEL

A seller sells a durable good at times $t = 1, 2, \dots, T$, where $T \leq \infty$. At each time t , a single buyer arrives and decides if and when to buy the object.¹² All parties discount the future at rate δ . The product is costless for the seller to produce,¹³ while each buyer has unit demand. Throughout what follows, we let t denote calendar time, and let a index a buyer's arrival time. Each buyer has an independently drawn discounted lifetime utility from purchasing the object. We let v denote some unspecified buyer's value, and assume that each buyer's value is drawn from a distribution F supported on \mathbb{R}_+ , with $0 < \mathbb{E}[v] < \infty$. We let \underline{v} denote the minimum value in the support of F . The distribution F is fixed and common knowledge, and buyer values for the object do not change over time.

However, the buyers do not directly know their v ; instead, they learn about it through signals they obtain over time, via some information structure. To be precise, a *dynamic information structure* \mathcal{I}_a for a buyer arriving at time a is:

- A set of possible signals for every time t after a , i.e., a sequence $(S_t)_{t=a}^T$, and
- Probability distributions given by $I_{a,t} : R_+ \times S_a^{t-1} \times P^t \rightarrow \Delta(S_t)$, 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. Note that the buyer observes signal realization s_t at time t , whose distribution depends on (their own) true value $v \in R_+$, the history of (their own) previous signal realizations $s_a^{t-1} = (s_a, s_{a+1}, \dots, s_{t-1}) \in S_a^{t-1}$, as well as the history of *all previous and current* prices $p^t = (p_1, p_2, \dots, p_t) \in P^t$. In particular, this definition allows for information structures to display history dependence.¹⁴

The timing of the model is as follows. At time 0, the seller commits to a pricing strategy σ , which is a distribution over possible price paths $p^T = (p_t)_{t=1}^T$. We allow $p_t = \infty$ to mean that the

¹²Our basic analysis is unchanged if the number of arriving buyers varies over time, or is stochastic, provided the value distribution is fixed.

¹³Introducing a cost of c per unit does not change our results: it is as if the value distribution F were "shifted down" by c , and the buyer might have a negative value. The transformed distribution G in Definition 1 below would also be shifted down by c .

¹⁴To avoid measurability issues, we assume each signal set S_t is at most countably infinite. All information structures in our analysis have this property.

seller refuses to sell in period t . Note that the price the seller posts at time t must be the same for all buyers that have the ability to buy in that period. After the seller chooses the strategy, nature chooses a dynamic information structure for each buyer. In each period $t \geq 1$, the price in that period p_t is realized according to $\sigma(p_t | p^{t-1})$. A buyer arriving at time a with true value v observes the signal s_t with probability $I_{a,t}(s_t | v, s_a^{t-1}, p^t)$ and decides whether or not to purchase the product (and if so, when).

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

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

The inner expectation $\mathbb{E}[v | s_a^{\tau}, p^{\tau}]$ 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 evolution of prices and signals. We note that the stopping time τ_a is allowed 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 structure chosen by nature were the worst possible, given his pricing strategy σ and buyer's optimizing behavior. Hence the seller's payoff is:

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

Note that when a buyer faces indifference, ties are broken against the seller. Breaking indifference in favor of the seller would not change our results, but would add cumbersome details.¹⁵

2.1. Discussion of Assumptions

Several of our assumptions are worth commenting on. First, following the robust mechanism design literature, we assume that the buyer has perfect knowledge of the information structure whereas the seller does not. More precisely, each buyer knows the information structure, and is Bayesian about what information will be received in the future. In contrast, the seller is uncertain about the information structure itself. Our interpretation is that the buyers understands what information they will have access to; for instance, someone may always rely upon some product review website and hence know very well how to interpret the reviews. The seller, on the other hand, knows that there are many possible ways buyers can learn, and wants to do well against all

¹⁵When ties are broken against the seller, it follows from our analysis that the sup inf is achieved as max min. This would not be true if ties were broken in favor of the seller.

these possibilities. In Section 9.1, we will show that our results extend even if the seller knows the buyer begins with extra prior information (say, through advertising). Thus, a deterministic constant price path remains optimal when nature is constrained to provide some particular information (but could provide more) in the first period. In Appendix D.2, we also show that, as long as the buyer is uncertainty averse and knows how to interpret all signals they have received, the worst case for the seller involves a Bayesian buyer. Our results only require that the buyers know what information they receive at any given time, and so our assumptions are actually not significantly more strict than what would arise in a single period model.

Second, we assume that the value distribution is common knowledge. This restriction is for simplicity, allowing us to focus on information arrival and learning. The assumption also enables us to compare our results to the classic literature on intertemporal pricing. In fact, the classic setting where the buyer knows her value can be seen as an extreme case of our extended model in Section 9.1, where the buyer has a more informative (non-degenerate) prior and may receive additional information over time.

Third, we assume that the information structure for a buyer arriving at a only depends on their value, the signal history for that buyer and the price history. In principle, one may want the information structure to depend on more variables, such as the purchasing history, or the signals and values of other buyers. However, because of our worst case objective and the IID assumption, allowing for nature to condition on more variables would not hurt the seller further. We comment on the implications of this assumption in Section 8.

Fourth, we assume that the seller commits to a pricing strategy that the buyer observes, and chosen before the information structure is determined. There are two components of this restriction. First, we assume the seller must have some method of committing to (and communicating to the buyer) a particular randomization in the future; although, as it turns out, randomization does not help the seller under our timing assumption. Studying robust intertemporal pricing with limited commitment is left to future work. Second, we *restrict* the seller to using pricing mechanisms, and rule out, for instance, mechanisms that randomly allocate the object as a function of reports. We view this as a restriction on the environment, albeit one that tends to be quite common in our applications. We comment that a result of Riley and Zeckhauser (1983) implies that, if buyers always knew their true values in the model above, these mechanisms would achieve the optimal second-best profit. This point suggests that the restriction should not, by itself, influence the interpretation of the results.

Finally, our key timing assumption is that the information structure in each period is determined *after* the price for that period has been realized. As in the literature review, if the information structure is determined *before* the price is realized, then the one-period optimal seller strategy in this model follows from Roesler and Szentos (2017) and Du (2017). The question of timing is more

delicate under dynamics; should a buyer’s second period information depend on the first period price they observed? What about buyers that arrive later? In Section 7, we consider a number of compelling alternative timing assumptions for which dynamic extensions of the robust selling mechanism of Du (2017) recover our results (or cannot). We emphasize that we do *not* believe we are at all in conflict with Du (2017), who focuses on settings where the seller has access to general mechanisms (and can thus randomize), nor Roesler and Szentes (2017), where a buyer chooses the information structure before the price is set.

Still, the idea that information could depend (at least somewhat) on price in practice seems intuitive. When shopping online, a consumer’s information about a product may depend on how prominently it is displayed in the search results. If the buyer sorts products by how expensive they are, then the information structure will depend on the realized price. Or, the seller may update his price at the beginning of every quarter, with buyers having the ability to purchase at that price in the next three months. What information (e.g. product reviews or competitors’ advertisements) buyers receive following each quarter’s price announcement may well depend on the realized price. Given these observations, it makes sense to start the dynamic analysis with the most cautious timing assumption, in order to avoid taking a stand on which restrictions are reasonable or not. This approach is the one the model above follows.

3. SINGLE PERIOD ANALYSIS

We start with the case where the seller does not worry about intertemporal incentives. For simplicity, we do this by taking $T = 1$, although the results are identical if buyers are myopic or could only purchase upon arrival. To solve this problem, we define a transformed distribution of F . For expositional simplicity, the following definition assumes F is continuous. All of our results in this paper extend to the discrete case, though the general definition requires additional care and is relegated to Appendix A.

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

The pressed distribution G is useful because for any (realized) price p , nature can only ensure that the object remains unsold with probability $G(p)$. This holds since the worst-case information structure has the property that a buyer who does not buy has expected value exactly p . To see why, consider an information structure where the buyer’s belief following a recommendation to not buy is $v_N < p$ and the recommendation following a recommendation to buy is $v_B(> p)$. Then

nature could, with some small probability $\varepsilon > 0$, give the recommendation to not buy whenever buy would have been recommended, hurting the seller. The buyer would have a higher belief following a recommendation to not buy, but would still follow it if ε were sufficiently small. This logic holds as long as $v_N < p$.

In fact, the worst case information structure following a price p is a *partition* with a threshold that induces a belief p following the recommendation not to buy. One can show (e.g., Kolotilin (2015)) that partitional information structures minimize the probability the buyer is recommended to buy, whenever the belief following the recommendation to not buy must be some fixed value (in our case, p). This observation allows us to show that the worst-case information structure involves telling the buyer whether her value is above or below $F^{-1}(G(p))$, making $1 - G(p)$ the probability of sale.

These remarks give us the following proposition:

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

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

For future reference, we call p^ the one-period maxmin price and similarly $\Pi^* = \max_p p(1 - G(p))$ the one-period maxmin profit.*

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

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

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

Finally, also note that there are other information structures which induce the same worst-case profit for the seller. For example, the buyer could be told her value exactly if it is above the threshold, since she will still buy. However, any worst case information structure involves the buyer being told if her value is *below* the threshold (i.e., the lowest element of the partition cannot be refined further on a set of positive measure).

4. INTERTEMPORAL INCENTIVES DO NOT HELP

In this section we present our first main result, that having multiple periods to sell does not allow the seller to extract more surplus from each buyer.¹⁶ Stokey (1979) demonstrated that this result holds when buyers know their values, provided they do not change over time. On the other hand, she also demonstrated that if values do change over time, letting buyers delay purchase could enable a seller to obtain higher profits by facilitating price discrimination.¹⁷ One may wonder whether information arrival, which affects the buyers' value over time, could similarly make price discrimination worthwhile.

However, it turns out for worst case information structures, these concerns do not arise. For simplicity, we focus on the case where there is a single buyer at time 1, since the argument readily extends to the case where buyers arrive at every time. With only the first buyer, the seller could always sell exclusively in the first period, the one-period profit Π^* forms a lower bound on the seller's maxmin profit from this buyer. To show that Π^* is also an upper bound, we explicitly construct a dynamic information structure for any pricing strategy, such that the seller's profit under this information structure decomposes into a convex combination of one-period profits. Our proof takes advantage of the partitioned form of worst-case information structures from the single period problem:

Proposition 2. *For any pricing strategy $\sigma \in \Delta(p^T)$, there is a dynamic information structure \mathcal{I} and a corresponding optimal stopping time τ^* that lead to expected (undiscounted, per-buyer) profit no more than Π^* . So, for a single buyer, the seller's maxmin profit against all dynamic information structures is Π^* , irrespective of the time horizon T and the discount factor δ .*

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

Let us first review the sorting argument when the buyer knows her value. In this case, given a price path $(p_t)_{t=1}^T$, we can find time periods $1 \leq t_1 < t_2 < \dots \leq T$ and value cutoffs

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

¹⁷It is interesting to note in our worst case information structures, buyers who do not buy actually *do* have a positive continuation value, even though this need not hold for arbitrary (non-worst case) partitioned information arrival processes.

$w_{t_1} > w_{t_2} > \dots \geq 0$, such that the buyer with $v \in [w_{t_j}, w_{t_{j-1}}]$ optimally buys in period t_j (see e.g. Stokey (1979)). This implies that in period t_j , the object is sold with probability $F(w_{t_{j-1}}) - F(w_{t_j})$.

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

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

This information structure is similar to the one period problem, in that a buyer is told whether her value is above a given threshold. But unlike the one period problem, if the buyer's value is below the threshold, she still has positive expected surplus from continuing. Instead, she is indifferent between purchasing *and continuing without further information*. We describe the buyer's optimal stopping behavior in the following lemma:

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

The proof of this lemma can be found in Appendix A, where we actually prove a more general result for random prices.

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

Proof of Proposition 2 for Deterministic Prices. Since the buyer with true value v in the percentile range $(G(w_{t_j}), G(w_{t_{j-1}})]$ buys in period t_j , the seller's discounted profit is given by (assuming $T = \infty$):

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

where the second line is by Abel summation,¹⁸ the third line is by w_{t_j} 's indifference between buying in period t_j or t_{j+1} , and the last inequality uses $w_{t_j}(1 - G(w_{t_j})) \leq \Pi^*$, $\forall j$. For finite horizon T , the proof proceeds along the same lines except for a minor modification to Abel summation. ■

Relative to the potential complexity of arbitrary information arrival processes, we find it noteworthy that the information structures constructed here are reasonably intuitive: Consumers buy when they find out that their value is above some (price contingent threshold). While the particular thresholds rely upon the worst-case objective, they are still relatively straightforward to find. Given an arbitrary price path, describing the information structure is no more difficult than computing the function G and finding the buyer indifference thresholds.

Despite the appeal of the analogy to the known value case, it is worth noting that for an arbitrary declining price path, these information structures *may not* be the worst a seller may face for a given declining price path. The following example illustrates this:

Example 2. Let $T = 2$, $v \in \{0, 1\}$ with $\mathbb{P}[v = 1] = 1/2$ and $\delta = 1/2$. Suppose the seller were to use a price path $p_1 = 11/40$ and $p_2 = 1/10$. Since a buyer would be indifferent between purchase and delay with a true value of $\frac{9}{20}$, the information structure constructed in Lemma 1 applied to this example induces posterior expected value $\frac{9}{20}$ when the buyer is recommended to not purchase in the first period, and expected value p_2 when recommended to not purchase in the second period. One can show that the (overall) expected profit is:¹⁹

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

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

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

In fact, equality holds. Under this information structure, the seller's profit is therefore $\delta p_2 \mathbb{P}[v = 1] = \frac{1}{40} = 0.025 < 0.0427$.

The discrete value space is used for simplicity. The important feature of the example is that in the second period, the buyer *strictly* prefers following the recommendation they are given

¹⁸Abel summation says that $\sum_{j \geq 1} a_j b_j = \sum_{j \geq 1} \left((a_j - a_{j+1}) \sum_{i=1}^j b_i \right)$ for any two sequences $\{a_j\}_{j=1}^{\infty}$, $\{b_j\}_{j=1}^{\infty}$ such that $a_j \rightarrow 0$ and $\sum_{i=1}^j b_i$ is bounded. We take $a_j = \delta^{t_j-1} p_{t_j}$ and $b_j = G(w_{t_{j-1}}) - G(w_{t_j})$.

¹⁹If the probability of being recommended to buy in period t is r_t , we have $\frac{1}{2} = r_1 + \frac{9}{20}(1 - r_1)$ and $\frac{9}{20} = r_2 + \frac{1}{10}(1 - r_2)$. These equations give $r_1 = \frac{1}{11}$ and $r_2 = \frac{7}{18}$.

to disobeying it. While this is not a feature of partitional information structures in the single period problem, such information structures could be used in order to induce delay in dynamic settings. Explicitly solving for these worst case information structures seems challenging, and is not necessary for our main result on the optimality of constant prices.

While most of the intuition for the general result is captured by the above argument, random prices introduce a technical difficulty in applying the sorting argument directly. Specifically, since the threshold values w_{t_j} depend on both the realized price and the distribution of future prices, they are in general random variables. More problematically, these thresholds may be non-monotonic if they are to be defined using the buyer's indifference condition. If such non-monotonicity occurs, we will not be able to express the seller's discounted profit as a convex sum of one-period profits, and the above proof will fail.

The intuition from the deterministic case still works when prices can be random, but we introduce some technical tools in order to recover the proof. The details are in Appendix A. Specifically, we modify the relevant indifference thresholds so that they are *forced* to be decreasing. To be precise, we define v_t to be the smallest value (in the known-value case) that is indifferent between buying in period t at price p_t and optimally stopping in the future, and *then* let $w_t = \min\{v_1, v_2, \dots, v_t\}$. We think of this as keeping track of the "binding" thresholds, above which all consumers have already bought. Using this modified definition for w_t , we can consider the same information structure as in the above proof and show that Proposition 2 continues to hold for random prices.

5. OPTIMALITY OF CONSTANT PRICES

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

Under known values, any arriving buyer facing a constant price path would buy immediately (if she were to buy at all), due to impatience. However, the promise of future information may induce the buyer to delay. Given an arbitrary information arrival process and constant price path, even if the buyer had positive expected surplus from buying upon arrival, they may prefer to wait just to make sure. Nevertheless, in the following lemma, we show that in the worst case, the seller eliminates the potential damage of delayed purchase by committing to never lower the price.

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

We present the intuition here and leave the formal proof to Appendix A. Let us fix a non-decreasing price path. For any dynamic information structure nature can choose, we consider an alternative information structure that simply gives a recommendation to the buyer to either buy or not in the first period. The probability of receiving each recommendation depends on when they would have bought in the original (dynamic) information structure. In other words, we push nature's recommendation to time 1. The proof shows that for non-decreasing prices, we can find a replacement such that the buyer still follows the recommendation of whether to buy. The replacement information structure gives the recommendation to not purchase in a way that imposes the same cost on the buyer that originally arose due to delay. This replacement has the property that the seller's profit is (weakly) decreased.²⁰ Since the seller receives at least Π^* under any information structure that releases information only in the first period, we obtain the lemma.

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

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 optimal, and it is uniquely optimal whenever the one-period maxmin price p^* is unique.*

Facing a constant price path, a worst-case dynamic information structure for each buyer simply involves giving each buyer the same information structure they would have obtained with only a single period. Buyers either purchase immediately or never, and hence intertemporal incentives do not matter. This completes our analysis of the baseline model.

6. QUALITY DIMENSION

In the baseline model, the purchasing decision is 0-1, so that the buyer faces a simple stopping problem. But even in our main applications of interest, the consumer's problem may be more complicated. For instance, Apple often produces new products with varying quality levels, and most consumers would obtain as much surplus from owning one as they would from owning two. Our machinery will be useful for studying this case, although we will need to import some additional tools from the literature on information design beyond what we have already utilized.

We enrich the baseline model by assuming that when the buyer purchases, he chooses between a high quality and a low quality object. As before, the buyer has a type $v \sim F$. A high quality

²⁰Specifically, it is unchanged against a constant price path and strictly decreased against an increasing price path.

object requires the seller to pay an additional cost k to produce, but gives a surplus of v to the buyer. On the other hand, a low quality object is free to produce, and gives the buyer a surplus of vq . At each time period, the seller offers a menu of prices for each quality level (again using the convention that an infinite price means not selling in that period), denoting the price of the low quality object by p_L and the price of the high quality object by p_H . We again assume consumers only obtain surplus from a single item, and that they leave after purchase. The model is otherwise identical.

6.1. One Period

As with the baseline model, we begin by analyzing the $T = 1$ case. To do this, consider an arbitrary information structure that results in a posterior distribution of values \tilde{F} . We can write the seller profits as:

$$\begin{aligned}\Pi &= (p_H - k)(1 - \tilde{F}(v_H)) + p_L(\tilde{F}(v_H) - \tilde{F}(v_L)) \\ &= ((1 - q)v_H - k)(1 - \tilde{F}(v_H)) + qv_L(1 - \tilde{F}(v_L))\end{aligned}\tag{3}$$

where $v_H = \frac{p_H - p_L}{1 - q}$ and $v_L = \frac{p_L}{q}$ are the threshold values at which a consumer would be indifferent between purchasing the high quality good and the low quality good, and the low quality good and nothing, respectively.

For this section, we will let Π^* denote the seller profit when only one quality level is offered, and let Π^{**} denote the seller profit when two quality levels are offered. We have the following:

Proposition 3. *Suppose k is sufficiently small and that $p^* > \underline{v}$ (recalling that p^* is the optimal single period price from the model without quality gradations). Then the seller gains from offering a menu of two quality levels; that is, $\Pi^{**} > \Pi^*$.*

The proof of this proposition utilizes a geometric representation of information structures, based on Rothschild and Stiglitz (1970) (elaborated on by Gentzkow and Kamenica (2016)):

Lemma 3 (Rothschild and Stiglitz (1970), Gentzkow and Kamenica (2016)). *If \tilde{F} is a distribution of posterior expected values, then*

$$\int_0^x \tilde{F}(s) ds \leq \int_0^x F(s) ds, \quad \forall 0 \leq x \leq 1.\tag{4}$$

This result is useful because it allows us to give a joint upper bound of $\tilde{F}(v_H)$ and $\tilde{F}(v_L)$. Our analysis from the one-period benchmark shows that $\tilde{F}(v_i) \leq G(v_i)$ (because $G(v_i)$ maximizes

the probability of no sale). By using (4), we can further show that for an appropriate choice of (v_H, v_L) , the inequality $\tilde{F}(v_i) \leq G(v_i)$ must be sufficiently slack at some v_i . Intuitively, if this were not the case, \tilde{F} would have to arise from a bipartite information structure with two different thresholds (corresponding to v_H and v_L). Together with (3), we are able to show $\Pi^{**} > \Pi^* + \varepsilon$. Details are left to the Appendix.

The geometric characterization described in Lemma 3 allows us to illustrate the worst case information structure (or, more precisely, the distribution of posterior values that arises from the worst case information structure):

Corollary 1. *The worst case information structure has support $\{v_L, v_H, v^*\}$, where v_L is the threshold value at which the buyer is indifferent between low quality and not purchasing, v_H is the threshold value at which the buyer is indifferent between high quality and low quality, and $v^* > v_H$. The signal inducing expected value v^* can be taken to be a partition.*

One aside on the one-period analysis: the reason the seller prefers to offer multiple quality levels is fairly straightforward in this model—doing so limits the ability of nature to utilize partitional information structures to hurt the agent’s surplus. That said, the result may also be viewed as noteworthy since at $k = 0$, the low quality object can be seen as a “damaged good,” and in fact we could even allow the low quality good to be the one that is costly to produce.²¹ Deneckere and McAfee (1996) showed that the conditions for damaged goods to be optimal in Bayesian models are surprisingly stringent given how prevalent they are in practice. Our result provides a novel (and comparatively elegant) justification for the provision of damaged goods: if consumers are unsure about their value, then damaged goods can lead to a better worst-case guarantee for the seller.

6.1.1. Single period optimal menu examples

These examples illustrate the optimal menu when $v \sim \text{Uniform}[0, 1]$, demonstrating that non-partitional information structures arise and illustrating some features of the optimum. For the following examples, we recall that when only high quality is offered, the resulting optimum is $\Pi^* = 1/8$ with a price $p^* = 1/4$, and consumers being recommended to purchase whenever $v > 1/2$ (and not purchase when $v < 1/2$). These computations utilize techniques from Gentzkow and Kamenica (2016) and Kolotilin, Li, Mylovanov and Zapechelnyuk (2017). The details are in the Appendix C:

²¹This is seen easily by observing that given $k = 0$, there is both a positive probability of both items being offered and strictly higher profits than with one quality level, meaning that if k is applied to the low quality item the seller still does strictly better by offering both.

Example 3. Let $v \sim U[0, 1]$, $k = 0$ and $q = 1/2$. Then the optimal menu involves the monopolist offering a menu with $p_L \approx 0.1210$ and $p_H \approx 0.3098$, corresponding to $v_L \approx 0.2419$ and $v_H \approx 0.3778$. The worst case information structure involves telling the buyer whether the value is above $\approx .5899$, and recommendation no purchase with probability $\approx .3596$. The seller's profit is therefore $\Pi^{**} \approx 0.1549$.

One noteworthy feature of this example is that $v_L(1 - \tilde{F}(v_L)) = v_H(1 - \tilde{F}(v_H))$. It turns out that this feature is special to the case of $q = 1/2$.

Example 4. Let $v \sim U[0, 1]$, $k = 0$ and $q = 2/3$. Then the optimal menu involves the monopolist offering a menu with $p_L \approx 0.1621$ and $p_H \approx 0.3036$, corresponding to $v_L = 0.2431$ and $v_H = 0.4246$. The worst case information structure involves the buyer being told whether the value is above ≈ 0.5831 and being recommended to not purchase with probability ≈ 0.4274 . The seller's profit is therefore $\Pi^{**} \approx 0.1518$, lower than the profit with $q = 1/2$.

We can verify that the profit is greater than $1/8$, although we point out that there is non-monotonicity of the optimal profit in q .

6.2. Dynamics

In this section, we suppose that the seller uses the same menu in every period and illustrate how intertemporal incentives arise when dynamics are added (in contrast to the baseline model). To provide intuition, we describe a simple information structure that induces delay:

- Reveal no information when the buyer arrives, and
- Endow the buyer with the single period worst case information structure in the period after the buyer's arrival.

This information structure will induce delay when δ is large enough, in contrast to the case with a single quality level. With a single quality level, the buyer is indifferent between purchasing and not when they do not purchase, and hence future information does not *improve* their decisions. In contrast, with differing quality levels, information can be used to induce delay since it can result in the buyer strictly increasing her payoffs by making a better decision on the quality dimension.

For the single quality level case, Lemma 2 replaces an arbitrary dynamic information structure with one that reveals all information to the buyer immediately upon arrival. While the signal the buyer obtains upon arrival is more informative under the replacement (since they observe a garbling of the future history), the choice of replacement ensures that the buyer is still willing to not purchase if given this recommendation. So this replacement may have made some purchasing buyers more optimistic, but as long as the purchasing probabilities were appropriately adjusted,

this would not matter. To similarly ensure that the probability of purchase adjusted appropriately, we would again make some buyers more optimistic. However, this could help the seller, by inducing buyers to make a different choice at the *intensive* margin. In the following proof, we construct an information structure where a buyer is indifferent between waiting and buying the low quality item. In the following period, however, they are provided information which makes them sufficiently optimistic so that they would purchase the high quality item. The increase in optimism overwhelms the impact of the discount factor, hurting the seller:

Claim 1. *Suppose the seller utilizes a constant menu equal to the single period optimum, which involves a non-partitional worst case information structure in the single period optimum. Then, for all δ , the seller's maxmin profit is less than $\Pi^{**} \cdot \frac{1-\delta^T}{1-\delta}$.*

The result that intertemporal incentives arise imply that calculating the optimal menu policy is significantly more difficult in general, but we are able to comment on optimal pricing strategies under some circumstances. For instance, if arrivals were significantly frontloaded, then a policy of phasing out products may dominate the constant menu strategy. Doing so prevents delay, which is a significant potential cost when arrivals are sufficiently frontloaded. In contrast, with a single quality level, refusing to sell a product at a later period can only hurt the seller in the worst case, since future profits are lost when this is done. Worst case information structures need not induce delay, and so a phase-out would never be preferred to the constant pricing strategy.

7. TIMING

This section analyzes the implications of our assumption regarding the timing of information acquisition relative to pricing. This assumption is captured in how we define dynamic information structures, since we allow them to be contingent on all past prices as well as the current price, but not future prices. When $T = 1$, since information is only provided once, in our view there are two natural benchmarks²² to be interested in: first, the case where information can depend on the (single) price, and second, the case where it cannot. The former is analyzed in Section 3; the latter is more difficult, but is studied in Roesler and Szentes (2017) and Du (2017), which together solve the seller's (and nature's) problem in this benchmark.

For completeness, we recall their result. To make the connection with our paper most clear, we impose as in these papers that the buyer's value distribution F is supported on $[0, 1]$. Roesler and Szentes (2017) observe that in choosing an information structure, nature is equivalently choosing

²²One could also study cases where information interacts somewhat, but not arbitrarily, with the price. We do not do so here since we are not aware of any reasonable, compelling alternative restrictions.

a distribution \tilde{F} of posterior expected values, such that F is a mean-preserving spread of \tilde{F} .²³ They solve for the worst-case distribution \tilde{F} as summarized below:

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

$$F_W^B(x) = \begin{cases} 0 & x \in [0, W) \\ 1 - \frac{W}{x} & x \in [W, B) \\ 1 & x \in [B, 1] \end{cases} \quad (5)$$

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

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

As alluded to in Section 2.1, the issue of defining dynamic information structures is more subtle than defining static information structures, for a variety of reasons that we hope this section will elucidate. Briefly, there are many more ways for information to interact with price (or not interact with the price) in the dynamic setting. Our benchmark corresponds to the most cautious case. However, other cases are worth commenting on as well, especially since it is natural to be interested in describing how Roesler and Szentes (2017) and Du (2017) extend to dynamic settings, and our main model does not do this.

We first consider re-defining a *dynamic information structure* to be a sequence of signal sets $(S_t)_{t=1}^T$ and probability distributions $I_{a,t} : R_+ \times S_a^{t-1} \times P^{t-1} \rightarrow \Delta(S_t)$. The crucial distinction from our main model is that the signal s_t depends on previous prices p^{t-1} but not on the current price p_t . The seller chooses a pricing strategy that achieves maxmin profit against such information structures and corresponding optimal stopping times of the buyer.

Before moving to the full model with arriving buyers, as a warm up we consider the case where there is only a single buyer but a longer horizon. In our model, the fact that the seller cannot do better with a longer horizon is demonstrated in Proposition 2. This relied upon the

²³This equivalence is separately observed by Gentzkow and Kamenica (2016) in the context of Bayesian persuasion. These authors attribute the result to Rothschild and Stiglitz (1970).

construction of a dynamic information structure that held the seller down to a profit of Π^* per buyer. But these information structures are not feasible in the Roesler and Szentes (2017) and Du (2017) setting. We have the following:

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

This proposition holds by the following reasoning (which is more direct than our Proposition 2): With multiple periods and a single buyer, the seller can guarantee Π_{RSD} by selling only once in the first period (using Du's mechanism). On the other hand, suppose nature provides the Roesler-Szentes information structure in the first period and no additional information in later periods. Then the seller faces a fixed distribution of values given by F_W^B . By Stokey (1979), selling only once is optimal against this distribution, and the seller's optimal profit is at most $W = \Pi_{RSD}$. This proves the result. In other words, Both Proposition 2 and Proposition 4 show that regardless of the timing of nature's moves, a longer selling horizon does not help the seller. Here, this conclusion follows from the duality between Roesler-Szentes and Du—as the above proof shows, Proposition 4 continues to hold even if nature only provides information in the first period. In our main model however, nature had to counter every pricing strategy with a *dynamic* information structure.

We now analyze three alternative cases of timing assumptions in the dynamic context.

7.1. Case One: Information *only* upon arrival, but possibly contingent on past (though not current) prices

Our first benchmark takes the definition of dynamic information structures from above, but assumes that each buyer is only endowed with a single signal. Hence each buyer is endowed with a single probability distribution $I_a : R_+ \times P^{t-1} \rightarrow \Delta(S_a)$. These are dynamic in the sense that they respond to prices, but not in the sense that information arrives over time.

In this setting, a constant price path cannot achieve $\Pi_{RSD} \cdot \frac{1-\delta^T}{1-\delta}$ whenever $\Pi^* < \Pi_{RSD}$. Indeed, nature can provide $F_{\Pi_{RSD}}^B$ to the first buyer, and the worst case partitional information structure to all subsequent buyers. Doing so delivers a lower profit than Π_{RSD} , which could be obtained by conditioning on the arrival time. However, the seller can change the way in which the price is randomized over time in order to deliver the discounted sum of single period profits:

Theorem 2. *Suppose there are arriving buyers, and suppose each buyer only receives information once upon arrival (before the price realizes in that period). For any time horizon T and any discount factor δ , the seller has a pricing strategy that ensures profit at least Π_{RSD} from each buyer. Thus the seller's maxmin profit is $\Pi_{RSD} \cdot \left(\frac{1-\delta^T}{1-\delta}\right)$.*

The proof is based on a key lemma (Lemma 6 in Appendix B) relating the outcome under a static price distribution to that under a dynamic price distribution. This outcome-equivalence property enables us to construct a dynamic pricing strategy that replicates Du’s mechanism for each arriving buyer, achieving Π_{RSD} as profit guarantee. Essentially, we consider randomization over *threshold indifference conditions* instead of randomization over price itself, and then use this to construct the dynamic random pricing strategy of the seller.

7.2. Case Two: Information cannot depend on prices at all

Next, we consider the case where we define a dynamic information structure to be *completely* price independent—that is, a probability distribution $I_{a,t} : R_+ \times S_a^{t-1} \rightarrow \Delta(S_t)$ —but allow for buyers to obtain information over time. While this case is in some sense the polar opposite of what we study, the resulting optimal is remarkably similar. Specifically, an optimal strategy for the seller is to utilize a constant price path, though drawing the price path randomly according to a Du distribution instead of setting it equal to p^* (as opposed to in our main model). It turns out that again, intertemporal incentives disappear when this strategy is employed:

Theorem 3. *Suppose there are arriving buyers, and suppose that all information is independent of all realized prices (though may depend on the pricing strategy). For any time horizon T and any discount factor δ , the seller has a pricing strategy that ensures profit at least Π_{RSD} from each buyer. Thus the seller’s maxmin profit is $\Pi_{RSD} \cdot \left(\frac{1-\delta^T}{1-\delta}\right)$.*

This theorem introduces some new techniques that may be of independent interest. Recall that, in allowing for random pricing strategies in Section 4, we defined cutoff values (forced to be decreasing) in two steps—first using the buyer’s indifference condition, and then keeping track of the lowest realized value arising from this indifference condition. Intuitively, these were the relevant “binding thresholds,” above which consumers would have already bought. Inspired by this technique, the proof of Theorem 3 introduces the dual definition of cutoff thresholds, namely *cutoff prices*. This approach is natural since for this proposition, we started with a sensible guess for the worst case pricing strategy (instead of attempting to construct an arbitrary information structure as in Section 4). While still requiring a bit of work, we are able to use the cutoff prices to show that the lower bound from an arbitrary dynamic information structure coincides with the benchmark without information arrival. The remaining details are left to the appendix.

7.3. Case Three: Information can depend upon past (though not current) prices and is dynamic

We conclude by adopting the definition of dynamic information structures from Case 1, but allow for information to be released to buyers over time. We have shown that when information can

be released over time, but not depend on the price, a constant price path can obtain Π_{RSD} from each buyer. When information cannot be released over time, but can depend on past prices, the seller can again obtain Π_{RSD} from each buyer, although not utilizing a constant price paths. This suggests that if we allow for both fully dynamic information arrival and partially price-contingent information, the seller may not be able to do as well. We show that this is indeed the case:

Claim 2. *Consider a model with two periods and one buyer arriving in each period. Suppose nature can provide information dynamically (to the first buyer). Assume that $\Pi_{RSD} > \Pi^*$ and that Du's mechanism is uniquely maxmin optimal in the one-period problem. Then the seller's maxmin profit in this two-period model with arriving buyers is strictly below $(1 + \delta)\Pi_{RSD}$ for any $\delta \in (0, 1)$.*

While the proof of this claim is fairly involved, the information structure chosen by nature is simple. When a buyer arrives, nature provides her with the Roesler-Szentes information structure. This yields profit at most Π_{RSD} from the second buyer, and similarly from the first buyer if she expects no additional information in the second period. We show that nature can induce delayed purchase from the first buyer and further damage profit by promising future information. Specifically, nature can reveal the value perfectly, in the second period, to any buyer who *would have* purchased in the first period without any additional information. The key technical step of the proof shows that delay always hurts the seller, and it occurs with strictly positive probability.²⁴

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

On the other hand, we have stated Claim 2 with an extra assumption that Du's mechanism is strictly optimal. This is for technical reasons that we explain in Appendix B, and it may not be necessary for the conclusion. In any event, we show this assumption holds for *generic* F .

8. INFORMATIONAL EXTERNALITIES

This section modifies the model from Section 2 to allow for interdependent preferences and information to be conveyed across buyers. Notice that *both* modifications must be made in order for the solution to the seller's problem to be altered. Any information generated by other

²⁴We are only able to show that for this specific information structure, total profit is strictly below $(1 + \delta)\Pi_{RSD}$. Since this is generally not the worst-case information structure for every pricing strategy, we do not know how to solve for the actual maxmin profit in the model considered here.

buyers is simply a restriction on nature's problem, and the worst case profit could only go up by introducing constraints. And such information is meaningless unless one buyer's value influences the conditional distribution of the other buyer's value. In contrast, we will show that when both features are present, then the seller will be able to achieve a higher expected discounted profit.

We replace the independent value assumption with the other extreme, where all buyers share the same value, assuming that $v \sim F$ is drawn at the beginning of the interaction. We also assume that nature chooses a single information arrival process, \mathcal{I} , consisting of signal sets $(S_t)_{1 \leq t \leq T}$ and distributions $I_t : R_+ \times S^{t-1} \times P^t \rightarrow \Delta(S_t)$ that are observed by all parties. In particular, there are two additional restrictions we are imposing on nature's choice of information arrival processes. First, a buyer that arrives at time t observes the signals observed by any buyer at time $t' \leq t$. Second, all buyers that have the opportunity to buy at time t observe the same signal realization. We return to this in Example 5.

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

Lemma 4. *Consider the model with common values and public signals suppose buyers are short lived. Fix a weakly increasing price path (p_1, \dots, p_T) with $p_1 \leq p_2 \leq \dots \leq p_T$. Then the worst case profit is achievable by an information structure that involves a single signal that is observed by all buyers.*

This result is the analog of Lemma 2 for this setting. The proof considers a similar replacement for an arbitrary dynamic information structures, resulting in an information structure consisting of a single signal.

Lemma 4 implies that for increasing price paths, it is sufficient to look at a single distribution of posterior valuations to find optimal price paths. We can write the seller's discounted profit from an increasing price path as:

$$\Pi^C = \min_{\tilde{F}} \sum_{t=1}^T p_t \delta^{t-1} (1 - \tilde{F}(p_t)), \quad (6)$$

where \tilde{F} is some distribution of posterior expected values arising from an information structure (i.e., satisfying the constraint (4)).

Using the similarity of (6) to the static model with quality (from Section 6), we observe that when $T = 2$, an increasing price path delivers higher profits than the constant price path via an identical argument as Proposition 3. Additionally, we can use the same information structure as in Lemma 1 to hold the seller down to a profit of Π^* per buyer for any weakly declining price path.

This reasoning immediately proves:

Corollary 2. *Take $T = 2$. Then with common values and public signals, whenever $p^* > \underline{v}$, the seller optimally obtains an average profit greater than Π^* from each buyer by utilizing introductory pricing (i.e., $p_1 < p_2$).*

For longer horizons, concluding that introductory pricing is optimal is more difficult (though it seems to be a reasonable conjecture). One can show weakly decreasing prices are suboptimal for general T via the same argument as for $T = 2$. However, the task remains to obtain a meaningful upper bound for the seller's maxmin optimal profit against non-monotonic pricing strategies. This was not a concern in the independent value setting, because information structures were constructed buyer-by-buyer. The single period profit does not always work when prices are non-monotonic, since increasing prices can deliver higher payoff to the seller. When information structures are correlated across buyers, a future price increase could influence the information provided to all buyers.

Despite this difficulty, we *are* able to show that there is a sequence of introductory pricing strategies that become optimal as $\delta \rightarrow 1$ when $T = \infty$:²⁵

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

$$\lim_{\delta \rightarrow 1} (1 - \delta)\Pi^C(\delta, \infty) = \Pi_{RSD} \quad (7)$$

which is achieved by a sequence of introductory (i.e., strictly increasing) price paths.

The proof of this theorem uses that an upper bound for the left hand side of (7) is obtained via a public Roesler-Szentes information structure for all buyers. We then use our (random price) Du mechanism to construct a sequence of price paths such that as $\delta \rightarrow 1$, the expression for the average per-period profit converges to the single period profit under a Du mechanism. These price paths, for uniformly distributed values, are shown in Figure 2, for $\delta = 9/10$ and $\delta = 95/100$ (fixing the initial price to be Π_{RSD}). We see that they involve the monopolist raising prices steeply at first, eventually leveling out. When the public signal is the Roesler-Szentes information structure, then the probability that the monopolist sells in every period is bounded away from 0, since the largest value in the support of the constructed price distribution is strictly below the supremum of the support in the Roesler-Szentes distribution (see the Appendix for details).

²⁵The same result could be obtained if we replaced the left hand side of (7) with $\lim_{T \rightarrow \infty} \frac{1}{T}\Pi^C(1, T)$

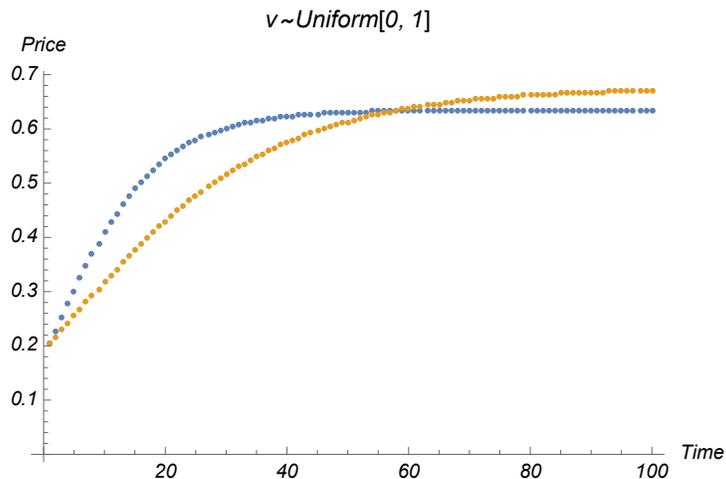


Figure 1: Illustration of constructed price paths. Blue is $\delta = 9/10$, orange is $\delta = 95/100$.

8.1. Short lived buyers versus long lived buyers

The problem of optimal pricing when information is conveyed across buyers has been studied in several other papers in other Bayesian settings, such as Bose et. al. (2006, 2008) (as well as papers cited therein). A major difference between this literature and our setting is that we allow buyers to delay purchase.²⁶ Partially motivated by this distinction, we now comment on a direct comparison between our benchmark and one in which buyers *cannot* delay, and must purchase upon arrival (which we dub the case of “short lived buyers”). Our main comment is that for increasing price paths, the profit levels coincide in the worst-case:

Proposition 5. *Consider the model with common values and public signals. For a price path (p_t) , let $\Pi_S((p_t))$ be the seller’s profit when buyers can only purchase upon arrival and $\Pi_L((p_t))$ be the seller’s profit when buyers can delay. Then whenever the price path $(p_t)_{t=1}^T$ is increasing, we have $\Pi_S((p_t)) = \Pi_L((p_t))$.*

The proof of this proposition shows that a worst case information structure is of the same form as in Lemma 4, although the argument is different since buyers can only buy upon arrival by construction. An immediate corollary is that the pricing strategies utilized in the proof of Theorem 4, inspired by the random price mechanisms from Section 7, also achieve approximate optimality in the cases where buyers cannot delay. We also obtain the same optimal introductory price paths as in the case where buyers can delay:

²⁶While some papers do allow for buyers to make decisions over time, as far as we are aware, these all involve “small” buyers who have a negligible impact on the information structure. This feature does not hold in our setting, since delay by any buyer would influence information arrival.

Corollary 2'. *Take $T = 2$. Then with common values and public signals, whenever $p^* > \underline{v}$ and buyers are short lived, the seller optimally obtains an average profit greater than Π^* from each buyer by utilizing introductory pricing (i.e., $p_1 < p_2$).*

This corollary implies that the maxmin optimal price paths are the same in the case where buyers can delay as when they cannot. We note that the proof differs slightly from the argument used to show Corollary 2, because a different information structure is used against a declining price path to hold the seller down to a profit of no more than Π^* per buyer. With short lived buyers, the partitional information structure induces belief p_1 in the first period, whereas it induces belief w_1 with long lived buyers.

Finally, recall that we assume that first, later buyers observe the signals of earlier buyers, and second, signals are publicly observed by all buyers with the opportunity to buy in the same period. The first restriction implies that nature is more constrained in providing information to later buyers. The second restriction eliminates the possibility for delay to hurt the seller's profits, due to our replacement Lemma. The following shows that dropping the second restriction allows nature to utilize information structures that induce delay in order to hurt the seller:

Example 5. *Take $T = 2$, and suppose the common value is $v \in \{0, 1\}$ with $\mathbb{P}[v = 1] = 1/2$ with $\delta = 1$ (the same conclusion will hold for δ sufficiently high). Suppose the seller utilizes prices $p_1 = 2/5$ and $p_2 = 1/2 > p_1$. With short lived buyers, if the first buyer observed the second buyer's signal following any delay, the worst case information structure would involve a signal observed by both buyers, whose posterior expected value is supported on $\{2/5, 1/2, 1\}$. Under this restriction, one can show that the worst case information structure induces posterior value equal to $2/5$ with probability $5/6$, and posterior value of 1 otherwise, yielding profit to the seller of $\frac{3}{20}$.*

We show that if the second period signal need not be public, the seller does strictly worse. First, suppose no information is provided to the first period buyer in the first period, and no information is provided to the second buyer in the second period. However, in the second period, the buyer from the first period obtains a signal such that the probability of having posterior expected value 1 in the second period is $1/5$. The first buyer is willing to delay, since purchasing in period 1 yields payoff $1/10$, whereas the payoff from delay is $(1/5) \cdot (1/2) = 1/10$. In this case, the seller's profit is $\frac{1}{10} < \frac{3}{20}$.

This example shows that the conclusion of Proposition 5 requires the restriction to public signal structures. Note that this issue of whether information is public *within the same period* only arises in cases where delay is possible, and is therefore not a distinction that arises when buyers are short lived. Example 5 emphasizes the importance of this distinction in general.

9. OUR ROBUSTNESS CONCEPT

9.1. Seller Initial Information

Our model so far assumes that the seller has no knowledge over the information the buyer receives. In practice, however, the seller may know that the buyer has access to at least some information. For example, he may conduct an advertising campaign, and understand its informational impact very well (Johnson and Myatt (2006)). While it may be impossible or difficult for such an advertising campaign to remove all uncertainty, the seller may nevertheless know that the buyer has access to some baseline information.²⁷ In this section we show that this possibility does not change our conclusions.

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

A signal s_0 induces a posterior belief on the buyer's value, which we denote by the distribution F_{s_0} . Define G_{s_0} to be the transformed distribution of F_{s_0} , following Definition 1. The same analysis as in Section 3 yields the following result:

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

$$p_{\mathcal{H}}^* \in \operatorname{argmax}_p p(1 - \mathbb{E}_{s_0}[G_{s_0}(p)]). \quad (8)$$

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

The expression (8) is familiar in two extreme cases: if \mathcal{H} is perfectly informative, then F_{s_0} is the point-mass distribution on s_0 . This means $G_{s_0}(p)$ is the indicator function for $p \geq s_0$, so that $\mathbb{E}_{s_0}[G_{s_0}(p)] = F(p)$. In contrast, if \mathcal{H} is completely uninformative, we return to Equation (1).

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

Proposition 2'. *In the multi-period model where the buyer observes initial information structure \mathcal{H} , the seller's maxmin profit against all dynamic information structures is $\Pi_{\mathcal{H}}^*$, irrespective of the time*

²⁷Note that if the seller has complete control over what information he provides, it would be impossible to do better than the full information outcome because nature could always reveal the value.

horizon T and the discount factor δ .

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

9.2. Information versus Taste

An equivalent (albeit more abstract) formulation of our model would be to assume that each buyer observes her value perfectly, but where the value follows a *stochastic process*, with the buyer instead drawing a vector of values $(v_1, \dots, v_T) \sim F$, where each coordinate denotes the value from purchasing the object at a given time. Nature's problem can then be thought of as choosing the stochastic process from some set of possible stochastic processes. Of course, *some* restrictions on the stochastic processes must be made to avoid a degenerate solution with the seller making zero profits. But certainly there are other ways of doing this; for instance, one could follow Carrasco et. al. (2017) and assume instead that the distribution of buyer values must have some fixed mean and support. We give an example of such a calculation in Appendix C, where we show that if the seller only knows $\mathbb{E}[v] = 1/2$ with the distribution having support $[0,1]$, the robustly optimal price is $\frac{1}{2}(2 - \sqrt{2}) > 1/4$ and profit is $\frac{3}{2} - \sqrt{2} < 1/8$.

We prefer our formulation of the single period model (i.e., nature choosing an information structure and not a value distribution) because it highlights the tractability associated with partitional information structures, which turns out to be useful in the dynamic analysis. Certainly other formulations may also yield similarly helpful single period solutions. However, restricting to information arrival process also influences how values evolve over time. In some sense, the restriction on how the value evolves is the more significant one that separates information arrival from taste shocks.

To see why, consider the following formulation with a single buyer that highlights the restrictions imposed by information arrival. For simplicity, we focus on the case of $T = 2$. Suppose the first period buyer's surplus from purchasing in the first period, v_1 , is distributed according to \tilde{F} , a distribution chosen by nature to minimize the seller's profits given the pricing strategy. Suppose, in addition, \tilde{F} must second order stochastically dominate the distribution F (as in our single period model). In contrast, the second period value is $v_2 \sim \tilde{F}(\cdot | v_1)$, where this random and need *only* respect the condition that $\mathbb{E}[v_2 | v_1] = v_1$. We do not impose other restrictions beyond this (in particular, we do not restrict the support of the buyer's second period value). The buyer that arrives in the second period faces the same problem as in the baseline model.

In this setting, the proof of Proposition 2 carries over without change, since any stochastic process for the buyer's value can be induced by the dynamic information structures we construct. But actually the seller can be hurt much worse when there are arriving buyers:

Claim 3. *Consider the two period model, with one buyer arriving in each period, and suppose each buyer’s initial value is distributed according to a distribution \tilde{F} that second order stochastically dominates F . Suppose the only restriction on the first period buyer’s value in the second period is that the expectation is the equal to the first period problem. Then the maxmin optimal profit for the seller when there are two buyers is Π^* . Hence the seller does not benefit at all from the presence of a second buyer.*

This case is extreme due to the lack of support restrictions for the second period value, although similar (but less sharp) results could be obtained under more stringent restrictions on the evolution of the buyer’s value process. We take the extreme case simply to highlight the significance of information arrival most clearly. When nature is not restricted by information arrival, non-purchasing buyers delay because their second period value may result in negative surplus from purchasing (which is similar to what occurs in the extension with quality and in Section 7.3). In particular, in this benchmark, we are not able to “push the recommendation to time 1” as we are in the benchmark model (Lemma 2). This is transparent in part due to how extreme the case is; there is a value process such the first period buyer’s purchase probability is approximately 0 as long as $\delta > 0$ against a constant, non-negative price path. But for information structures, we cannot have $\mathbb{E}[v \mid \text{Don't buy}] \leq p$ and $\mathbb{P}[\text{Don't buy}] \approx 1$ whenever $p < \mathbb{E}[v] < \infty$.²⁸

For general models with taste shocks, it may not make sense to restrict the mean in the second period to be equal to the mean in the first period, as done in this section. The restriction was simply made here in order to relate the model where the value evolution is restricted by information arrival to one where it is not. Finding sensible restrictions on the evolution of values under taste shocks that avoids degenerate solutions (or more general conditions that would yield our same results) is left to future work.

10. CONCLUSION

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

²⁸This is seen by recalling that for a given price, the probability a buyer does not purchase is maximized by a partitional information structure which tells the buyer whether or not the value is below $F^{-1}(G(p))$. Whenever $p < \mathbb{E}[v] < \infty$, there must be some positive probability that v is above this threshold.

introductory pricing can be optimal, and intertemporal incentives arise against a constant menu when there are gradations in quality.

Certainly our baseline setting is one that can (and we believe should) be studied in other contexts and under other modeling assumption. The case of limited seller commitment seems compelling, though there are technical difficulties similar to those that arise when formalizing (seller) learning under ambiguity (see Epstein and Schneider (2007)). One could also ask similar questions in more general dynamic mechanism design settings, where the agent's problem may not be represented by the choice of a stopping time. And settings with competition, changing values, richer population dynamics and different seller objectives all seem intriguing as well.

This paper contributes to a growing literature which employs the maxmin approach in analyzing the optimal design of mechanisms. The literature has mostly focused on static settings, although we suspect dynamic settings will receive significant attention in the future. It turned out the maxmin objective was useful in two respects:

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

These results allowed us to circumvent the difficulties that would otherwise be associated with profit maximization in arbitrary models of information arrival. We hope our analysis has suggested ways that such models could be analyzed to produce new economic insights. While it is certainly worthwhile to analyze other dynamic models, doing so would first require a way of handling information arrival in a tractable way, similar to what we have done here.

References

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

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

Dirk Bergemann, Benjamin Brooks, and Stephen Morris. Informationally robust optimal auction design. Technical report, Working Paper, Yale University, University of Chicago and Princeton University, December 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, January 2017.

- Dirk Bergemann and Karl Schlag. Robust monopoly pricing. *Journal of Economic Theory*, 146:2527–2543, 2011.
- Andrew Biehl. Durable-goods monopoly with stochastic values. *Rand Journal of Economics*, 32(3):565–577, Autumn 2001.
- Simon Board. Durable-goods monopoly with varying demand. *Review of Economic Studies*, 75(2):391–413, April 2008.
- Subir Bose, Gerhard Orosel, Marco Ottaviani, and Lise Vesterlund. Dynamic monopoly pricing and herding. *RAND Journal of Economics*, 37(4):910–928, Winter 2006.
- Subir Bose, Gerhard Orosel, Marco Ottaviani, and Lise Vesterlund. Monopoly pricing in the binary herding model. *Economic Theory*, 37: 203–241, 2008.
- Jeremy Bulow. Durable-goods monopolists. *Journal of Political Economy*, 90(2):314–332, April 1982.
- Rene Caldentey, Ying Liu, and Ilan Lobel. Intertemporal pricing under minimax regret. *Operations Research*, November 2016.
- Vinicius Carrasco, Vitor Farinha Luz, Nenad Kos, Matthias Messner, Paulo Monteiro, and Humberto Moreira. Optimal selling mechanisms under moment conditions. Technical report, Working Paper, PUC-Rio, April 2017.
- Gabriel Carroll. Robustness and linear contracts. *American Economic Review*, 105(2):536–563, February 2015.
- Gabriel Carroll. Robustness and separation in multidimensional screening. *Econometrica*, 85(2):453–488, March 2017.
- Sylvain Chassang. Calibrated incentive contracts. *Econometrica*, 81(5):1935–1971, September 2013.
- Yiwei Chen and Vivek Farias. Robust dynamic pricing with strategic customers. Technical report, Working Paper, Renmin University and MIT, December 2016.
- Kim-Sau Chung and Jeffrey Ely. Foundations of dominant-strategy mechanisms. *Review of Economic Studies*, 74:447–476, 2007.
- John Conlisk, Eitan Gerstner, and Joel Sobel. Cyclic pricing by a durable goods monopolist. *Quarterly Journal of Economics*, 99(3):489–505, August 1984.

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

Songzi Du. Robust mechanisms under common valuation. Technical report, Working Paper, Simon Fraser University, January 2017.

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

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

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

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

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

Matthew Gentzkow and Emir Kamenica. A rothschild-stiglitz approach to bayesian persuasion. *American Economic Review: Papers and Proceedings*, 106(5):597–601, May 2016.

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

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

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

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

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

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

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

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

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

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

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

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

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

A. PROOFS FOR THE MAIN MODEL

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

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

Let \underline{v} be the minimum value in the support of F . For $\beta \in (\underline{v}, \mathbb{E}[v]]$, define $G(\beta) = \sup\{\alpha : g(\alpha) \leq \beta\}$. We extend the domain of this inverse function to \mathbb{R}_+ by setting $G(\beta) = 0$ for $\beta \leq \underline{v}$ and $G(\beta) = 1$ for $\beta > \mathbb{E}[v]$.²⁹

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

²⁹If 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.1. Proof of Proposition 1

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

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

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

A.2. Proof of Proposition 2

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

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

$$v - p_t \geq \max_{\tau \geq t+1} \mathbb{E}[\delta^{\tau-t} \cdot (v - p_\tau)] \quad (9)$$

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 $v \geq v_t$.³⁰ Thus, v_t is defined by the equation

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

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

³⁰This follows by observing that both sides of the inequality are strictly increasing in v , but the LHS increases faster.

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

$$w_t = \min\{v_1, v_2, \dots, v_t\} = \min\{w_{t-1}, v_t\}. \quad (11)$$

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

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

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

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

Proof of Lemma 1'. Suppose period t is the first time that the buyer learns her value is above the $G(w_t)$ -percentile. Then in particular, $w_t < w_{t-1}$ which implies $w_t = v_t$ by (11). Given this signal, she knows that she will receive no more information in the future (because w_t decreases over time). She also knows that her value is above the $G(w_t)$ -percentile, which is greater than $w_t = v_t$, the average value below that percentile. Thus from the definition of v_t , the buyer optimally buys in period t .

On the other hand, suppose that in some period t the buyer learns her value is below the $G(w_t)$ -percentile. Since w_t decreases over time, this signal is Blackwell sufficient for all previous signals. By definition of G , the buyer's expected value is $w_t \leq v_t$. Thus even without additional information in the future, this buyer prefers to delay her purchase. The promise of future information does not change the result. ■

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

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

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

Lemma 5. *Suppose $w_t = v_t \leq w_{t-1}$ in some period t . Then*

$$p_t = \mathbb{E} \left[\sum_{s=t}^{T-1} (1 - \delta) \delta^{s-t} w_s + \delta^{T-t} w_T \mid p^t \right] \quad (12)$$

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

Using Lemma 5, we can rewrite the profit as

$$\begin{aligned}
\Pi &= \mathbb{E} \left[\sum_{t=1}^T \delta^{t-1} \cdot (G(w_{t-1}) - G(w_t)) \cdot \mathbb{E} \left[\sum_{s=t}^{T-1} (1 - \delta) \delta^{s-t} w_s + \delta^{T-t} w_T \mid p^t \right] \right] \\
&= \mathbb{E} \left[\sum_{t=1}^T \delta^{t-1} \cdot (G(w_{t-1}) - G(w_t)) \cdot \left(\sum_{s=t}^{T-1} (1 - \delta) \delta^{s-t} w_s + \delta^{T-t} w_T \right) \right] \\
&= \mathbb{E} \left[\sum_{s=1}^{T-1} (1 - \delta) \delta^{s-1} w_s (1 - G(w_s)) + \delta^{T-1} w_T (1 - G(w_T)) \right] \\
&\leq \Pi^*.
\end{aligned} \tag{13}$$

The second line is by the law of iterated expectations, because w_{t-1} and w_t only depend on the realized prices p^t . The next line follows from interchanging the order of summation, and the last inequality is because $w_s(1 - G(w_s)) \leq \Pi^*$ holds for every w_s . Hence it only remains to prove Lemma 5.

Proof of Lemma 5. We assume that T is finite. The infinite-horizon result follows from an approximation by finite horizons and the Monotone Convergence Theorem, whose details we omit. We prove by induction on $T - t$, where the base case $t = T$ follows from $w_T = v_T = p_T$. For $t < T$, from (10) 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

$$p_t = \mathbb{E}[(1 - \delta^{\tau-t})v_t + \delta^{\tau-t}p_\tau]. \tag{14}$$

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

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

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

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

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

Plugging the above two expressions into (14) proves the lemma. ■

A.3. Proof of Lemma 2

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

As discussed in the main text, we will construct another information structure \mathcal{I}' which only reveals information in the first period, and which weakly reduces the seller's profit. Consider a signal set $S = \{\bar{s}, \underline{s}\}$, corresponding to the recommendation of “buy” and “not buy”, respectively. To specify the distribution of these signals conditional on v , let nature draw signals s_1, s_2, \dots according to the original information structure \mathcal{I} (and conditional on v). If, along this sequence of realized signals, the stopping time τ results in buying the object, let the buyer receive the signal \bar{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 \bar{s} or \underline{s} in the first period and provides no more information afterwards.

We claim that under \mathcal{I}' , the buyer receiving the signal \underline{s} has expected value at most p_1 . We actually show something stronger, namely that the buyer has expected value at most p_1 conditional on the signal \underline{s} and *any* realized signal s_1 .³¹ To prove this, note that since stopping at time τ is weakly better than stopping at time 1, we have

$$\mathbb{E}[v \mid s_1] - p_1 \leq \mathbb{E}^{s_2, \dots, s_T} \left[\delta^{\tau-1} (\mathbb{E}[v \mid s_1, s_2, \dots, s_\tau] - p_\tau) \right]. \quad (15)$$

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, \dots, 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 (15) to the following.

$$\mathbb{E}[v \mid s_1] \leq \mathbb{E}^{s_2, \dots, s_T} \left[\delta^{\tau-1} \mathbb{E}[v \mid s_1, s_2, \dots, s_\tau] + (1 - \delta^{\tau-1})p_1 \right]. \quad (16)$$

³¹Technically we only consider those s_1 such that \underline{s} occurs with positive probability given s_1 .

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

$$p_1 \geq \frac{\mathbb{E}^{s_2, \dots, s_T} [(1 - \delta^{\tau-1}) \cdot \mathbb{E}[v \mid s_1, s_2, \dots, s_\tau]]}{\mathbb{E}^{s_2, \dots, s_T} [1 - \delta^{\tau-1}]} \quad (17)$$

The denominator $\mathbb{E}^{s_2, \dots, s_T} [1 - \delta^{\tau-1}]$ can be rewritten as $\mathbb{E}^{s_2, \dots, s_T} [\mathbb{P}(\underline{s} \mid s_1, s_2, \dots, s_T)]$, which is the probability of \underline{s} given s_1 . Because τ is a stopping time, the numerator in (17) can be rewritten as

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

which can be further rewritten as

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

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

With these, (17) states that

$$p_1 \geq \frac{\mathbb{E}^{s_2, \dots, s_T} [\mathbb{P}(\underline{s} \mid s_1, s_2, \dots, s_T) \cdot \mathbb{E}[v \mid s_1, s_2, \dots, s_T, \underline{s}]]}{\mathbb{E}^{s_2, \dots, s_T} [\mathbb{P}(\underline{s} \mid s_1, s_2, \dots, s_T)]} = \mathbb{E}[v \mid s_1, \underline{s}] \quad (18)$$

just as we claimed.

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

A.4. Proof of Theorem 1

By the previous lemma, a constant price path p^* delivers expected un-discounted profit Π^* from each arriving buyer. This matches the upper bound given by Proposition 2 and shows that always charging p^* is optimal. Moreover, suppose p^* is unique, then from (13) we see that the seller's profit from the first buyer equals Π^* only if $w_s = p^*$ almost surely. This together with Lemma 5 implies $p_1 = p^*$ almost surely. Analogous argument for later buyer shows that the seller must always charge p^* to achieve the maxmin profit. Hence the proposition.

A.5. Uncertainty Leads to Lower Price

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

Proposition 6. *For any continuous distribution F , let \hat{p} be an optimal monopoly price under known values:*

$$\hat{p} \in \operatorname{argmax}_p p(1 - F(p)). \quad (19)$$

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

Proof of Proposition 6. 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,

$$pG(p) = \int_0^{F^{-1}(G(p))} v dF(v), \forall p \in [\underline{v}, \mathbb{E}[v]]. \quad (20)$$

Differentiating both sides with respect to p , we obtain

$$G(p) + pG'(p) = \frac{\partial}{\partial p}(F^{-1}(G(p))) \cdot F^{-1}(G(p)) \cdot F'(F^{-1}(G(p))) = G'(p) \cdot F^{-1}(G(p)). \quad (21)$$

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

$$G'(p) = \frac{G(p)}{F^{-1}(G(p)) - p}. \quad (22)$$

Thus,

$$G(p) + pG'(p) = \frac{G(p) \cdot F^{-1}(G(p))}{F^{-1}(G(p)) - p}. \quad (23)$$

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

$$\frac{p}{1-G(p)} \cdot \left(1 - F\left(\frac{p}{1-G(p)}\right)\right) < p. \quad (24)$$

From the definition of \hat{p} , we see that the LHS above is at most $\hat{p}(1 - F(\hat{p})) \leq \hat{p} < p$, as we claim to show. Moreover, when $\hat{p} > \underline{v}$, the last inequality $\hat{p}(1 - F(\hat{p})) < \hat{p}$ is strict. Tracing back the

previous arguments, we see that $G(p) + pG'(p) > 1$ holds even at $p = \hat{p}$. In that case we would have the strict inequality $p^* < \hat{p}$ as desired. ■

B. PROOFS FOR THE ALTERNATIVE TIMING MODEL

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

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

Proposition 7. $\Pi_{RSD} \geq \Pi^*$ with equality if and only if $W = \underline{v}$ ($= p^*$), where W is as defined in the Roesler-Szentes information structure (5). Furthermore, as the distribution F varies, the ratio Π_{RSD}/Π^* is unbounded.

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

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

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

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

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

Because the distribution F has two-point support, it is clear that nature can induce any \tilde{F} supported on $[0, 1]$ with mean λ as the distribution of posterior expected values. Thus the Roesler-

Szentes information structure involves the smallest W such that F_W^B has mean λ for some $B \leq 1$. From (5), we compute that the mean of F_W^B is $W \log B - W \log W + W$. We look for the smallest W such that $\log B = \frac{\lambda}{W} + \log W - 1$ is non-positive. It follows that W is the smallest positive root of the equation

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

For λ small, we have the approximation $\Pi_{RSD} = W \approx \frac{\lambda}{|\log \lambda|}$. Thus both ratios $\hat{\Pi}/\Pi_{RSD}$ and Π_{RSD}/Π^* are unbounded.³² ■

B.2. Proof of Theorem 2

Throughout, we represent the robust selling mechanism in Du (2017) by a random price, with c.d.f. $D(x)$; the details of this distribution can be found later in (35), but they are not relevant for this proof. Because nature can provide each arriving buyer with the Roesler-Szentes information structure (5), the seller at most obtains Π_{RSD} from each buyer. To complete the proof, we will construct a dynamic pricing strategy that yields Π_{RSD} from each buyer.

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

Lemma 6. *Fix any continuous distribution function D , any horizon T and any discount factor $\delta \in (0, 1)$. There exists a distribution of prices $\sigma \in \Delta(p^T)$ such that if a buyer arrives in period t and knows her value to be v , then her discounted probability of purchasing the object (discounted to period t) is equal to $D(v)$.*

In words, for any static pricing strategy there is a dynamic pricing strategy which does not condition on buyers' arrival times, but which results in the same outcome as the static prices for every type of each arriving buyer.

We state the lemma for continuous distributions so that the buyer's optimal stopping time is almost surely unique. From Du (2017), Du's distribution D is continuous except when it is a point-mass on W . In the latter case $\Pi_{RSD} = \Pi^*$, and Theorem 2 follows from Theorem 1.

Lemma 6 is useful for our problem because it implies, via the Revenue Equivalence Theorem, that a seller using strategy σ obtains the same profit from any buyer as if he sells only once to this buyer at a random price distributed according to D . This is true whenever the buyer's value distribution is determined upon arrival and fixed over time, which is what we assume for the current proposition. Since Du's static mechanism guarantees profit Π_{RSD} from every buyer, the proposition will follow once we prove the lemma.

³²We conjecture that these profit ratios become bounded under certain regularity conditions on F .

Proof of Lemma 6. We will first prove the result for $T = 2$, then generalize to all finite T and lastly discuss $T = \infty$.

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

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

$$v_1 - p_1 = \delta \cdot \mathbb{E}^{p_2 \sim D} [\max\{v_1 - p_2, 0\}]. \quad (25)$$

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

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

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

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

Suppose H is the c.d.f. of v_1 . Since $w_1 = v_1$ and $w_2 = \min\{v_1, p_2\}$, the probability that w is greater than x is given by $(1 - \delta)(1 - H(x)) + \delta(1 - H(x))(1 - D(x))$.³³ This has to be equal to $1 - D(x)$, which implies

$$1 - H(x) = \frac{1 - D(x)}{1 - \delta D(x)}. \quad (26)$$

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

Before proceeding, we remark that (26) implies the distribution H has the same support as D . However, (25) suggests that when v_1 achieves the maximum of this support, p_1 is in general strictly smaller than v_1 (unless the support is a singleton point, a case we have discussed). Intuitively, charging this maximum price in the first period leads to delayed purchase by buyers with high values, which is costly for the seller. On the other hand, the minimum price p_1 is indeed equal to the minimum of the support of D , which we denote by W ; when D is Du's distribution, this is

³³ $1 - H(x)$ is the probability that $w_1 > x$, and $(1 - H(x))(1 - D(x))$ is the probability that $w_2 > x$.

the same W as in the Roesler-Szentes information structure (5).

Step 2: Extension to finite T . We conjecture a pricing strategy σ that is independent across periods: $d\sigma(p_1, \dots, p_T) = d\sigma_1(p_1) \times \dots \times d\sigma_T(p_T)$, where we interpret each σ_t as a distribution. Define the cutoff values v_1, \dots, v_T as in (10). Note that due to independence, v_t only depends on current price p_t but not on previous prices.

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

The result of the lemma requires each $w^{(t)}$ to be distributed according to D . Simple calculation shows this is the case if $v_T \sim D$ and $v_1, \dots, v_{T-1} \sim H$ (since v_t depends only on p_t , they are independent random variables).³⁴ We can then solve for the price distributions $\sigma_1, \dots, \sigma_T$ by backward induction: σ_T must be D , and once the prices in period $t+1, \dots, T$ are determined, there is a one-to-one relation between p_t and v_t by (10). Thus, the distribution of p_t is uniquely pinned down by the desired distribution of v_t .

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

$$v_1 - p_1 = \max_{\tau \geq 2} \mathbb{E} [\delta^{\tau-1} (v_1 - p_\tau)]. \quad (27)$$

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 $v_1 - p_1$. (27) thus reduces to

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

which can be further simplified to

$$v_1 = p_1 + \frac{\delta}{1-\delta} \cdot \mathbb{E}^{p_2} [\max\{p_1 - p_2, 0\}]. \quad (28)$$

Let $P(x)$ denote the c.d.f. of p_1 (and of p_2). When $p_1 = x$, (28) implies

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

³⁴The reason $H(x)$ should be the c.d.f. of v_1 is best understood in the infinite horizon problem (see below). Under stationarity, the buyer with value x buys in period t with probability $H(x)$, conditional on not buying previously. Thus the discounted allocation probability is $\sum_t \delta^{t-1} (1-H(x))^{t-1} H(x)$. Setting this equal to $D(x)$ yields (26).

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

$$P(x) = H\left(x + \frac{\delta}{1-\delta} \int_0^x P(z) dz\right). \quad (29)$$

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

$$Q(x) = x + \frac{\delta}{1-\delta} \int_0^x P(z) dz; \quad U(y) = 1 + \frac{\delta}{1-\delta} H(y) = \frac{1}{1-\delta D(y)}. \quad (30)$$

(29) is the differential equation

$$U(Q(x)) = Q'(x). \quad (31)$$

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

$$\frac{\partial V(Q(x))}{\partial x} = V'(Q(x)) \cdot Q'(x) = \frac{Q'(x)}{U(Q(x))} = 1. \quad (32)$$

Inspired by the analysis for finite T , we conjecture that the minimum value of p_1 is W . That is, we conjecture $Q(W) = W$. Since $V(W) = W$, we deduce from (32) that $Q(x)$ is characterized by

$$V(Q(x)) = x \text{ with } V(y) = \int_0^y (1 - \delta D(z)) dz. \quad (33)$$

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

$$P(x) = \frac{1-\delta}{\delta} \cdot (Q'(x) - 1). \quad (34)$$

Lemma 6 is proved, and so is Theorem 2. ■

B.3. Proof of Theorem 3

Consider a constant price p randomly drawn according to the Du distribution. Recall that p is supported on $[S, W]$, and its density is $\frac{1}{\log \frac{S}{W}} \cdot \frac{1}{p}$. We show the seller's discounted expected profit is at least W .

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

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

and then

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

This is exactly dual to the definition of cutoff values, and whenever $q_t = r_t \geq q_{t-1}$, we have (See Lemma 5):

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

If the random price p satisfies $q_{t-1} \leq p < q_t$, then purchase occurs in period t . Total profit is thus:

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

where we define $\pi(y) = \min\{(y - W)^+, S - W\}$ to be the integral of $\log \frac{S}{W} \cdot p dD(p)$, and use the convention that $\pi(q_0) = 0$. In other words, $\pi(y) = 0$ for $y \leq W$, $\pi(y) = y - W$ for $y \in [W, S]$ and $\pi(y) = S - W$ for $y \geq S$. Define:

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

to be a modified version of the function π .³⁵

Indeed, $\hat{\pi}$ is smaller than π and strictly so when $y \leq W$. Then we have

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

³⁵Intuitively, π coincides with $\hat{\pi}$ if the threshold prices never fall below W . We cannot assume this a priori, although it is natural to expect that doing so would not be worst case, just as inducing a belief lower than the price to a non-purchasing buyer is not worst case in the main model.

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

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

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

The inequality holds since if γ is finite, then $v_\gamma - S \leq \mathbb{E}[(v_\infty - S)^+ \mid v_1, \dots, v_\gamma]$ by convexity. And if γ is infinite, then $\delta^{\gamma-1}(v_\gamma - S) = 0 \leq (v_\infty - S)^+$.

To summarize, we first showed $\log \frac{S}{W} \cdot \Pi \geq \mathbb{E} [(1 - \delta)\delta^{t-1}\hat{\pi}(q_t)]$, and since $\hat{\pi}$ is *concave*, this is smaller than $\hat{\pi}(\mathbb{E}[(1 - \delta)\delta^{t-1}q_t]) = \hat{\pi}(v_0)$. However, the lower bound $\mathbb{E}[\hat{\pi}(v_\infty)]$ holds by concavity, essentially because the distribution of v_∞ is more spread out than the cutoff prices q_t .

Letting \tilde{F} denoting the distribution of v_∞ , we have:³⁶

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

The first inequality follows from F being a mean preserving spread of \tilde{F} , and the second follows from (5). Hence $\Pi \geq W$ as desired. ■

³⁶The stronger result $\log \frac{S}{W} \cdot \Pi \geq \mathbb{E}[\pi(v_\infty)]$ would mean that profit is minimized by revealing all information at once, which would easily complete the proof. But in order to use concavity, we have had to work with $\hat{\pi}$ rather than π .

B.4. Proof of Claim 2

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

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

$$D(x) = \begin{cases} 0 & x \in [0, W) \\ \frac{\log \frac{x}{W}}{\log \frac{S}{W}} & x \in [W, S) \\ 1 & x \in [S, 1] \end{cases} \quad (35)$$

Here $S \in (W, B]$ is characterized by³⁷

$$\int_0^S F_W^B(v) dv = \int_0^S F(v) dv \quad (36)$$

where F_W^B is the Roesler-Szentes worst-case information structure (5). To explain further, Roesler and Szentes (2017) observe that the LHS in (36) must not exceed the RHS (for all S) because F is a mean-preserving spread of F_W^B . However, when W is smallest possible, this constraint must bind at some S .

The following observations will be crucial. Since the constraint $\int_0^x F_W^B(v) dv \leq \int_0^x F(v) dv$ binds at $x = S$, the first order condition gives $F_W^B(S) = F(S)$. This implies that not only F is a mean-preserving spread of F_W^B , but in fact the truncated distribution of F conditional on $v \leq S$ is also a mean-preserving spread of the corresponding truncation of F_W^B . In other words, the Roesler-Szentes information structure has the property that a buyer with true value $v \leq S$ only receives signal $\leq S$ (i.e., her posterior expected value is at most S), while a buyer with true value $v > S$ expects her value to be greater than S .

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

³⁷ S is strictly greater than W because otherwise D is a mass-point at W and $\Pi_{RSD} = \Pi^*$, contradicting the assumption of the proposition.

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

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

The penultimate equality uses (36) and the last one uses (5).

We note that in general, there could be more than one point S for which (36) holds. Thus, the maxmin optimal mechanism in one period need not be unique even if we restrict attention to the class of exponential mechanisms considered by Du (2017). But we do have the following result:

Lemma 7. *There is a unique maxmin optimal mechanism in the one-period simultaneous-move model if and only if (36) holds at a unique point S .*

We mention that for generic distributions F , there is a unique S that satisfies (36). However, the proof is tangential to the paper and we will leave it out. A sufficient condition is that $F(x)$ is convex, for example when F is uniform.³⁸

Proof of Lemma 7. “Only if” is obvious, so we focus on the “if” direction. Suppose S is unique, we need to show any random price that guarantees W must follow Du's distribution D . Suppose $r(p)$ is the p.d.f. of the random price, then profit is

$$\Pi = \int_0^1 p \cdot r(p) \cdot (1 - \tilde{F}(p)) dp. \quad (37)$$

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

By Roesler and Szentes (2017), $\tilde{F} = F_W^B$ is a solution to nature's problem. For this solution, the integral inequality constraint only binds at $x = S$. Standard perturbation techniques in the calculus of variations thus imply that $\tilde{F} = F_W^B$ cannot be improved upon only if $p \cdot r(p)$ is a constant for $p \in (W, S)$.³⁹ Similarly, $p \cdot r(p)$ must also be a constant on the interval $p \in (S, B)$; in fact, we can show this constant is zero.⁴⁰

³⁸Recall that $F(S) = F_W^B(S)$. However, $F(x) - F_W^B(x) = F(x) + \frac{W}{x} - 1$ is convex, and so it has at most two roots $x_0 < x_1$. Because $F(x) > F_W^B(x)$ for $x < x_0$, S being x_0 would contradict (36). Thus $S = x_1$ is unique.

³⁹Suppose to the contrary that $p \cdot r(p) > p' \cdot r(p')$ for some $p, p' \in (W, S)$. Then starting with $\tilde{F} = F_W^B$, nature could increase \tilde{F} around p and correspondingly decrease it around p' . The perturbed distribution \tilde{F} still satisfies the feasibility constraints, and the profit Π is reduced.

⁴⁰If this constant were $c > 0$, then on the interval $[S, B]$ nature seeks to minimize $c \cdot \int_S^B (1 - \tilde{F}(v)) dv$ subject to the integral inequality constraint and equal means: $\int_S^1 (1 - \tilde{F}(v)) dv = \int_S^1 (1 - F(v)) dv$. Thus nature equivalently

Hence, $r(p)$ must be supported on $[W, S]$ and $p \cdot r(p)$ is a constant. This condition together with $\int_W^S r(p) dp = 1$ uniquely pins down $r(p)$, which must be the density function associated with D . ■

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

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

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

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

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

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

Step 3: Buyer behavior and seller profit. To facilitate the discussion, we consider another information structure \mathcal{I}' in which nature reveals \tilde{v} in the first period but does nothing in the second period. Under \mathcal{I}' , the buyer's value distribution F_W^B does not change over time. Thus by Stokey (1979), the seller's profit would at most be Π_{RSD} . We will show that the seller's profit under the *dynamic* information structure \mathcal{I} could only be lower than under \mathcal{I}' (for any pricing strategy), and we also characterize when the comparison is strict.

There are three possibilities: first, if the price p_1 is relatively high so that $\tilde{v} < v_1(p_1)$, then the buyer does not buy in the first period under \mathcal{I}' . This is also her optimal decision under \mathcal{I} , because she will not receive extra information in the second period. Secondly, if the price is very low, then under both \mathcal{I} and \mathcal{I}' the buyer buys in the first period. Lastly, for some intermediate prices

maximizes $\int_B^1 (1 - \tilde{F}(v)) dv$. Choosing $\tilde{F} = F_W^B$ results in 0 and is sub-optimal.

the buyer buys in the first period under \mathcal{I}' but not under \mathcal{I} ; the opposite situation cannot occur because \mathcal{I} provides more information than \mathcal{I}' in the second period, and the buyer's incentive to wait could only be stronger.

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

To summarize, we have shown:

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

Step 4: Proof of the claim for σ^D . Let σ^D be the pricing strategy given by Lemma 6, which we recall is robust to information that arrives only once (for each buyer). Here we show that under the dynamic information structure \mathcal{I} , the seller's profit from the first buyer is strictly less than Π_{RSD} .

Recall from the proof of Lemma 6 that under σ^D , the price in the second period p_2 is drawn from Du's distribution D , independent of p_1 . On the other hand, p_1 is (continuously supported) on a smaller interval $[W, S_1]$, with $W < S_1 < S$; more precisely, the distribution of p_1 is determined by the condition that $v_1(p_1) \sim H$ (see (26)).

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

Now recall from Step 1 that a signal $\tilde{v} < S$ is only received when the true value also satisfies $v < S$. Because we assume $\Pi_{RSD} > \Pi^*$, Proposition 7 gives $W > \underline{v}$. Thus a positive measure of signals $\tilde{v} \in (W, S)$ is received when the true value v belongs to the interval $[\underline{v}, W]$.⁴¹ We claim that for any such \tilde{v} , knowing the true value in the second period strictly benefits the buyer. This is because according to her expected value $\tilde{v} > W$, the buyer in the second period buys at some price p_2 ; but if she were informed that $v < W$, she would not buy at any price p_2 (which is at least

⁴¹If $\Pi_{RSD} = \Pi^*$, then Proposition 7 implies $W = \underline{v} = p^*$ and Du's distribution is a mass-point at W . Information in the second period is irrelevant, because a buyer waiting till the second period always buys at price $p_2 = W = \underline{v}$.

W). This proves that by providing \mathcal{I} rather than \mathcal{I}' , nature induces a positive probability of delay. By Lemma 8, we deduce that profit from the first buyer is less than Π_{RSD} .

Step 5: Proof for an arbitrary pricing strategy σ . Finally, we turn to prove the proposition in its full generality. The argument is as follows (omitting technical details): suppose for contradiction that some pricing strategy σ guarantees profit almost Π_{RSD} from each buyer. Then because D is uniquely optimal in the one-period problem, the distribution of p_2 conditional on p_1 is “close” to D (in the Prokhorov metric) with high probability; otherwise nature could sufficiently damage the seller’s profit from the second buyer. Next, we can similarly show that the distribution of $v_1(p_1)$ under σ is close to H , which is its distribution if $\sigma = \sigma^D$.⁴² The rest of the proof proceeds as in Step 4: a positive measure of signals $\tilde{v} \in (W, S)$ is received when the true value satisfies $v < W$. For such \tilde{v} , full information in the second period is strictly valuable, and the buyer delays purchase if $v_1(p_1)$ is smaller than but close to \tilde{v} . By what we have shown, this occurs with strictly positive probability. But then Lemma 8 implies profit from the first buyer is bounded away from Π_{RSD} under \mathcal{I} , leading to a contradiction. The proof of Claim 2 is complete. ■

Let us conclude by commenting on the assumption that Du’s mechanism is uniquely optimal. Suppose this assumption fails, so that another point $\hat{S} > S$ satisfies (36). This means there are two different Du distributions D and \hat{D} , supported on $[W, S]$ and $[W, \hat{S}]$ respectively. On their supports, both of these distributions have density proportional to $\frac{1}{p}$ (see (35)). This observation allows us to write

$$\hat{D} = \alpha D + (1 - \alpha)E \quad (38)$$

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

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

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

⁴²Consider nature choosing \tilde{F} in the first period and doing nothing afterwards. The seller’s profit from the first buyer can be written as $\mathbb{E}^w[w(1 - \tilde{F}(w))]$, where the random variable w equals $v_1(p_1)$ with probability $1 - \delta$ and it equals $\min\{v_1(p_1), p_2\}$ with probability δ (see (13)). The distribution of w must be close to D , otherwise nature could choose \tilde{F} and damage profit from the first buyer. Since p_2 is approximately distributed according to D , we can derive as in the proof of Lemma 6 that $v_1(p_1)$ must be approximately distributed according to H .

Because the price in the second period follows a Du distribution, the seller's profit from the second buyer is at least Π_{RSD} . For the first buyer, consider first the information structure \mathcal{I}' as in Step 3, where nature reveals $\tilde{v} \sim F_W^B$ in the first period and no additional information afterwards. As shown in Footnote 42, the seller's profit from this buyer is $\mathbb{E}^w[w(1 - F_W^B(w))]$. This is as in the one-period model, where the seller charges price w and nature provides the Roesler-Szentes information structure.

Recall that w is a random variable that equals $v_1(p_1)$ with probability $1 - \delta$ and $\min\{v_1(p_1), p_2\}$ with complementary probability. Because $v_1(p_1) \sim E$, whose support is strictly above the support of p_2 , we deduce that $w \sim \delta D + (1 - \delta)E$. Thus, by (38), the distribution of w is a convex combination of D and \hat{D} whenever $\delta \geq \alpha$. Since the seller ensures profit Π_{RSD} by using a random price distributed according to either D or \hat{D} , he does just as well by charging w . We have thus shown that profit from the first buyer is at least Π_{RSD} under information structure \mathcal{I}' .

Moreover, we claim that when nature provides \mathcal{I} rather than \mathcal{I}' , no buyer delays her purchase. To see this, consider a buyer who purchases in the first period under \mathcal{I}' . By definition of v_1 and the fact that $v_1 \sim E$, this means the buyer's signal \tilde{v} in the first period satisfies $\tilde{v} \geq v_1(p_1) \geq S$. But then her true value v must also be at least S , as we showed in Step 1. Such a buyer purchases at any price $p_2 \in [W, S]$ regardless of any information in the second period. Thus, although nature promises future information under \mathcal{I} , this information does not improve the buyer's expected utility in the second period. Consequently the buyer's behavior under \mathcal{I} is the same as under \mathcal{I}' , and profit under \mathcal{I} is also equal to Π_{RSD} .

To summarize, we have constructed a pricing strategy σ such that if nature chooses the particular information structure \mathcal{I} (for the first buyer), the seller's total profit is at least $(1 + \delta)\Pi_{RSD}$. This explains why *our proof of Claim 2* requires the assumption that Du's mechanism is unique. We do not know whether the result generally holds without this assumption.⁴³

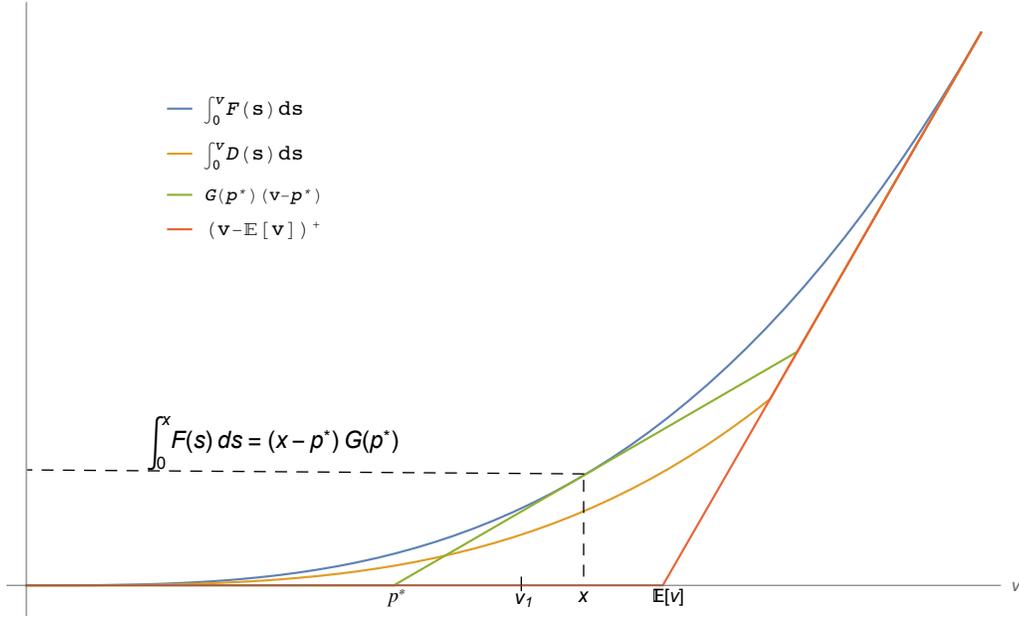
C. PROOFS FOR OTHER EXTENSIONS

Proof of Proposition 3. Note that there exists some $\bar{k} > 0$ such that if $k < \bar{k}$, the optimal menu when the seller is restricted to offering only *one* quality level involves high quality being sold (and this occurring with positive probability). We assume throughout that $k < \bar{k}$. In this case, the single period profit is equal to $(p^* - k)(1 - G(p^*))$, since the information structure the buyer faces does not depend on k .

We will take $v_L = p^*$ and $v_H > p^*$ to be determined.⁴⁴ From (4) and some feasible distribution

⁴³In other words, suppose S is not unique and suppose the seller uses the strategy σ constructed just now. We do not know whether nature can damage the seller's profit to be strictly lower than Π_{RSD} by choosing an information structure different from the \mathcal{I} in our proof.

⁴⁴Note that since $v_L = \frac{p_L}{q}$, we have $p_L = qp^* < p^*$.

Figure 2: Graphical illustration of the choice of v_1

of posterior values D , for all $x > v_H > p^*$ it holds that

$$\int_0^x F(s)ds \geq \int_0^x D(s)ds \geq (v_H - p^*)D(p^*) + (x - v_H)D(v_H), \quad (39)$$

where the second inequality follows from the monotonicity of D .

Let us choose x to be the buyer value corresponding to the $G(p^*)$ -percentile under F , i.e. $x = F^{-1}(G(p^*))$. Recall from the one-period problem that the lowest $G(p^*)$ -percentile of the distribution F has expected value p^* , so that (by $F(x) = G(p^*)$)

$$p^* = \frac{1}{F(x)} \int_0^x s dF(s). \quad (40)$$

Since $p^* > \underline{v}$, we have $G(p^*) > 0$ and $x > p^*$. Using integration by parts as well as the previous equation, we obtain

$$\int_0^x F(s)ds = xF(x) - \int_0^x s dF(s) = xF(x) - p^*F(x) = (x - p^*)G(p^*). \quad (41)$$

This step of our argument is illustrated in Figure 1, for the distribution $F(x) = x^2$.

Combining (39) and (41), we deduce the following inequality:

$$D(v_H) - G(p^*) \leq \frac{v_H - p^*}{x - v_H} \cdot (G(p^*) - D(p^*)). \quad (42)$$

By taking v_H sufficiently close to p^* , we can ensure that $\frac{v_H - p^*}{x - v_H} \leq \frac{q}{1 - q}$. Since $G(p^*) \geq D(p^*)$, we infer that

$$D(v_H) - G(p^*) \leq \frac{q}{1 - q} \cdot (G(p^*) - D(p^*)). \quad (43)$$

which we can rearrange to be:

$$1 - G(p^*) \leq (1 - q)(1 - D(v_H)) + q(1 - D(p^*)) \quad (44)$$

Note that (42) also implies $D(v_H) < 1$ for v_H sufficiently close to p^* , using $G(p^*) < 1$. This observation and the expression (3) lets us bound profit from below:

$$\begin{aligned} \Pi &= ((1 - q)v_H - k)(1 - D(v_H)) + qv_L(1 - D(v_L)) \\ &> (p^* - k)((1 - q)(1 - D(v_H)) + q(1 - D(p^*))) + \varepsilon + qk(1 - D(p^*)) \end{aligned} \quad (45)$$

where the inequality holds strictly for some ε sufficiently small. We emphasize that the choice of ε depends on $(1 - q)(v_H - p^*)(1 - D(v_H))$, but not on k directly. But by (44), we conclude that:

$$\Pi \geq (p^* - k)(1 - G(p^*)) + \varepsilon + qk(1 - D(p^*)) \geq \Pi^* + \varepsilon$$

for all $k \geq 0$. Since this holds for any such pricing strategy, we have $\Pi^{**} > \Pi^* + \varepsilon$. ■

Proof of Corollary 1. Note that the worst case information structure, under the assumptions of the previous proposition, has the buyer purchase both quality levels with positive probability. Via the usual revelation principle argument, it is without loss to consider information structures with 3 signals in the support.

Consider any distribution $D(v)$ that arises from some such information structure. Note that $\int_0^x D(s)ds$ must be tangent to $\int_0^x F(s)ds$ at some point $\bar{v} > v_H$; by our previous proof, otherwise, nature could choose an information structure that lowers the probability that the buyer purchases the highest quality object, hurting profits. It follows that any worst case information structure has the property that, if the buyer purchases low quality or does not purchase, then $v < \bar{v}$. Hence v^* is induced by a partition. On the other hand, given any threshold, the signal s_0 that maximizes the probability that the buyer does not buy satisfies $\mathbb{E}[v \mid s_0] = v_L$. Furthermore, if $\mathbb{E}[v \mid s_L] < v_H$, then we can also find some information structure with lower probability of sale that satisfies

$\mathbb{E}[v \mid s_L] = v_H$, as desired. ■

Lemma 9. *Whenever the single period problem has a non-partitional worst case information structure, there exist a continuum of worst case information structures which induce beliefs v_L and v_H . In particular, there exist information structures such that, for all v sufficiently close to (but below) the threshold inducing v^* , belief v_L is induced for a type of value v with positive probability.*

Proof of Lemma 9. Consider a worst case information structure inducing beliefs v_L, v_H and v^* , and suppose the signal inducing v^* is a partition with threshold \tilde{v} .

Fix some ε , and first consider the information structure that tells the agent if $v \in [\underline{v}, \underline{v} + \varepsilon]$, $v \in [\underline{v} + \varepsilon, \tilde{v} - \varepsilon]$, or $v \in [\tilde{v} - \varepsilon, \tilde{v}]$. There exists $\bar{\varepsilon}$ such that, whenever $\varepsilon < \bar{\varepsilon}$, there exists a distribution supported on v_L, v_H that arises from an information structure when $v \in [\underline{v} + \varepsilon, \tilde{v} - \varepsilon]$. Call this information structure \mathcal{I}_M . Additionally, we can find p and q such that if $v \in [\underline{v}, \underline{v} + \varepsilon]$, s_L is sent with probability p and s_H is sent with probability $1 - p$, and if $v \in [\tilde{v} - \varepsilon, \tilde{v}]$, s_L is sent with probability q and s_H is sent with probability $1 - q$, such that the expected value of v following a signal of s_L is v_L and the expected value of v following s_H is v_H . Call this information structure \mathcal{I}_O .

Consider the following information structure:

- The signal space is $\{v_L, v_H, v^*\}$.
- If $v > v^*$, the agent observes \tilde{v} (and purchases high quality).
- If $v \in [\bar{v} + \varepsilon, v^* - \varepsilon]$, the agent observes v_L or v_H according to \mathcal{I}_M .
- Otherwise, the agent observes v_L or v_H according to \mathcal{I}_O .

If v^* is induced, then $v > \tilde{v}$, which occurs with the same probability as under the original information structure. The only other two signals are induced when $v < \tilde{v}$. Furthermore, there are two possible signals on this event, which pins down the probabilities that each occur. Hence these also occur with the same probability as the original information structure, since any distribution with two point support is pinned down by the martingale condition. This observation completes the proof. ■

Details of quality and informational externalities examples: For these examples, we consider case where the seller's profit is:

$$\Pi_D = \alpha v_1(1 - D(v_1)) + (1 - \alpha)v_2(1 - D(v_2)),$$

with $\alpha v_1 < (1 - \alpha)v_2$ and $v_1 < v_2$. To move between examples, we can either adjust α , or multiply the expression by some constant k , noting that the expression will be maximized at the same value for v_1 and v_2 , and consider the resulting expression.

We first find the worst case D given v_1 and v_2 . Fixing $D(v_1)$, we seek a $D(v_2)$ that makes the previous expression as low as possible. By Corollary 1, the worst case $D(v_2)$ corresponds to a line that passes through $(v_2, D(v_1)(v_2 - v_1))$ which is tangent to the line $\int_0^x s ds = \frac{x^2}{2}$. The point of tangency occurs at an x solving:

$$\frac{x^2}{2} = x(x - v_2) + D(v_1)(v_2 - v_1).$$

Indeed, the right hand side is the line corresponding to the distribution D under consideration, where we observe that the slope of that line is equal to x at x . The left hand side is the line corresponding to complete information.

There are two roots to the above equation, with the positive root corresponding to the relevant solution. Since the value of this point is equal to the slope of the line under consideration, which in turn corresponds to $D(v_2)$, we have:

$$D(v_2) = v_2 + \sqrt{v_2^2 - 2D(v_1)(v_2 - v_1)}.$$

The worst case $D(v_1)$ therefore minimizes:

$$\alpha v_1(1 - D(v_1)) + (1 - \alpha)v_2(1 - (v_2 + \sqrt{v_2^2 - 2D(v_1)(v_2 - v_1)})).$$

The derivative of this expression with respect to $D(v_1)$ is:

$$\frac{(v_2 - v_1)v_2(1 - \alpha)}{\sqrt{v_2^2 - 2D(v_1)(v_2 - v_1)}} - v_1, \alpha$$

and the second derivative is positive whenever $v_2^2 > 2D(v_1)(v_2 - v_1)$, the same condition for some signal $D(v_2)$ to be used (noting that $D(v_1) \leq 2v_1$ for the information structure to be feasible). Therefore, whenever we have an interior solution:

$$D(v_1) = \left(\frac{v_2}{\alpha v_1}\right)^2 \frac{(\alpha v_2 - (v_2 - v_1))(\alpha v_1 + (1 - \alpha)(v_2 - v_1))}{2(v_2 - v_1)}$$

and

$$D(v_2) = v_2 + \frac{(v_2 - v_1)v_2(1 - \alpha)}{\alpha v_1}$$

This yields an expression for the seller's profit in terms of v_1 and v_2 (provided we are away from interior solutions).

The resulting expressions are somewhat unweildy, but can be used to numerically calculate the solution to the seller's problem. For $\alpha = 1/2$, after tedious algebra, one can show that the optimal choice of prices for the seller are:

$$v_1 = \frac{7 - \sqrt{17}}{49 - 9\sqrt{17}} \approx 0.2419, v_2 = \frac{3 + 11\sqrt{17}}{128} \approx .3778,$$

with

$$D(v_1) = \frac{1}{8}(7 - \sqrt{17}) \approx .3596, D(v_2) = \frac{1}{32}(23 - \sqrt{17}) \approx .5899.$$

Proof of Claim 1. Let \mathcal{I}_1 denote a worst case information structure in the single period problem, with signals "HIGH", "LOW" and "WAIT" corresponding to the buyer's actions. In particular, we assume that "WAIT" is given with positive probability to any $v < v^*$ whenever v is in some interval with endpoint v^* (which we can guarantee by Lemma 9). We consider the following set of dynamic information structures, parameterized by ε . For the first period:

- A signal is drawn according to \mathcal{I}_1 .
- Signals of "WAIT" and "LOW" are revealed to the buyer.
- Signals of "HIGH" are revealed if $v \geq v^* + \varepsilon$.
- If $v^* + \varepsilon > v \geq v^*$, "WAIT" is observed.
- In the second period, the buyer observes her true value if told to wait in the first period, and if $v^* - \alpha < v < v^* + \varepsilon$.

We choose α so that a buyer told to wait is indifferent between waiting and not. That is, we have (for ε small)

$$\int_{v^*}^{v^* + \varepsilon} (qs - p_L)f(s)ds = \delta \cdot \left(\int_{v^* - \alpha}^{v^* + \varepsilon} (s - p_H)f(s)ds \right). \quad (46)$$

Note in particular that, for any δ and a fixed α (sufficiently small so that all types with $v > v^* - \alpha$ would buy high, the left hand side vanishes as $\varepsilon \rightarrow 0$ but the right hand side does not. Hence this can always be satisfied.

That this signal is obedient follows from the obedience in the single period problem, together with the choice of α that makes the buyer indifferent between waiting and purchasing. We

compute the change in the seller's profit. To show that the seller is made worse off, it suffices to show that:

$$(1 - \delta)p_H(F(v^* + \varepsilon) - F(v^*)) > \delta p_H(F(v^*) - F(v^* - \alpha)),$$

or:

$$\int_{v^*}^{v^* + \varepsilon} f(s)ds > \delta \int_{v^* - \alpha}^{v^* + \varepsilon} f(s)ds. \quad (47)$$

However, since (46) holds, we infer that (47) would follow from the observation that:

$$qs - p_L < s' - p_H$$

for all $s \in [v^*, v^* + \varepsilon]$ and $s' \in [v^* - \alpha, v^* + \varepsilon]$. This clearly holds for ε small (since ε small ensures α is small as well), completing the proof. Hence some dynamic information structure can damage the seller's profits. ■

The first proofs for Section 8 gives us the same conclusion as Lemma 10 for the case of short lived buyers. After demonstrating the result for short lived buyers, we then move to long lived buyers subsequently:

Lemma 10. *Consider the model with common values and public signals suppose buyers are short lived. Fix a weakly increasing price path (p_1, \dots, p_T) with $p_1 \leq p_2 \leq \dots \leq p_T$. Then the worst case profit is achievable by an information structure that involves a single signal that is observed by all buyers.*

Proof of Lemma 10. Fix an increasing price path, and let $(\mathcal{I}_a)_{a=1}^T$ denote an arbitrary choice by nature for information structures for each arriving buyer at time a . First, without loss of generality, we can take the signal space of each \mathcal{I}_a to have two elements, following a recommendation to buy or not buy for each buyer. In addition to the conclusion of the lemma, we will show that any worst case information structure induces a buyer belief lower than p_1 with probability 0. The conclusion of the lemma holds for the case of $T = 1$ immediately, and this additional statement holds by Proposition 1.

Hence suppose we have shown that the conclusion holds for some horizon of length T , and now consider the case where there are $T + 1$ buyers. First, suppose the first buyer obtains a signal s_1 after which a purchase is not made. It follows that $\mathbb{E}[v \mid s_1] \leq p_1$, which, since the price path is increasing, implies $\mathbb{E}[v \mid s_1] \leq p_2$. It follows that the worst case information structure will not reveal any further information, since this holds the seller to a profit of 0. In this event, all buyers therefore can be taken to observe an identical signal s_1 .

On the other hand, suppose the first buyer purchases, so that $\mathbb{E}[v \mid s_1] > p_1$. Let the posterior distribution over the v that this induces be denoted \tilde{F}_{s_1} . By the inductive step, the worst case

information structure provides a single signal \tilde{s} to all buyers, and furthermore, that $\mathbb{E}[v \mid s_1, s] \geq p_2 \geq p_1$. Now consider an alternative information structure that reveals (s_1, \tilde{s}) to the first buyer. Then we still have, $\mathbb{E}[v \mid s_1, \tilde{s}] \geq p_2 \geq p_1$, meaning that even after observing the signal \tilde{s} , the first buyer is still willing to purchase. Hence in this case as well, we can assume all buyers observe the same signal. The claim then holds for any finite T by induction. The case of $T = \infty$ follows via an approximation argument. ■

Proof of Corollary 2'. Note that the seller's objective function is identical to equation (3) with $q = \delta \frac{1-\delta}{1-\delta^2}$ and $k = 0$. Proposition 3 shows that the seller does strictly better with an increasing price path, showing the proposition holds. Hence it follows that the seller can obtain more than Π^* from each buyer. Hence it suffices to find an information arrival process such that the seller obtains profits less than Π^* per buyer. Consider the following information structure:

- The first buyer is told if $v > F^{-1}(G(p_1))$
- The second buyer is told if $v > F^{-1}(G(p_2))$.

In particular, note that this information arrival process is feasible because $p_1 \geq p_2$. On the other hand, the surplus obtained from the buyer arriving at time t is $p_t(1 - G(p_t)) \leq \Pi^*$. Since there exists an increasing price path delivering more higher payoff per buyer, it follows that the optimal price path is increasing. ■

Proof of Lemma 4. We follow similar steps as in the proof of Lemma 10, proving the theorem by induction and then considering the infinite horizon case via finite approximation. The case of $T = 1$ is identical, so fix some arbitrary finite T . Consider an arbitrary dynamic information structure for the first buyer \mathcal{I} , which we assume contains at least as much information as is generated by the information structure after $t \geq 2$ (i.e., this buyer has the ability to see all information provided to later buyers). By the inductive step, we can take this information structure to be some time s (possibly equal to $T + 1$) at which point consumers stop purchasing, which all buyers arriving at time $t \geq 2$ observe (but the first buyer may not observe).

Following Lemma 2, replace the first buyer's information structure with $\{\underline{s}, 2, \dots, T + 1\}$, giving the recommendation \underline{s} if the information structure dictates that this buyer not buy, as well as with probability $1 - \delta^{\tau-1}$ if the information structure results in purchase at time τ . If the first buyer does not buy, then reveal no information to later buyers. If the first buyer does buy, then also reveal to this buyer the time s at which consumers stop purchasing. Note that this is a public information structure.

First, note that the payoff from the first buyer does not increase, again following the proof of Lemma 2. Since $\mathbb{E}[\delta^{\tau-1}p_\tau] > \mathbb{E}[\delta^{\tau-1}]p_1$ since the price path is increasing, we have that the profit from the first buyer is no larger. Furthermore, if the first buyer does not buy, then neither do any subsequent buyers, meaning that the seller's surplus is also minimized in this continuation history. On the other hand, if the first buyer does buy, then the original information structure revealed some time $s \geq 2$ at which point buyers stopped purchasing. Note that the any worst case information structure, beliefs are never lower than $p_s \geq p_1$, since otherwise a recommendation to not buy could be induced with higher probability. Hence the first buyer is still willing to follow the recommendation in the replacement. Therefore, the seller's profit can only decrease in the replacement information structure. ■

Proof of Theorem 4. Note first that $\lim_{\delta \rightarrow 1} (1 - \delta)\Pi^C \leq \Pi_{RSD}$, since nature can choose F_{RS} and reveal it publicly at time 1. In general, when nature chooses distribution D , we write the seller's average profit as:

$$\Pi^C((p_t)_{t=1}^\infty) = \sum_{t=1}^{\infty} p_t(1 - D(p_t))\delta^{t-1}(1 - \delta).$$

Let \tilde{F} denote the random price distribution (with density \tilde{f}) corresponding to Du's mechanism (See (35)). Let $(p_t^\delta)_{t=1}^\infty$ denote a sequence of price paths such that:

$$\tilde{F}(p_t^\delta) - \tilde{F}(p_{t-1}^\delta) \rightarrow \delta^{t-1}(1 - \delta)$$

with $p_1 = \Pi_{RSD} = W$. For any such price path, we have:

$$\Pi^C((p_t^\delta)_{t=1}^\infty) \rightarrow \int_W^B p(1 - D(p))\tilde{f}(p)dp = \Pi_{RSD},$$

as claimed. ■

Example from Section 9.2. Suppose $v \sim \tilde{G}$, where \tilde{G} is some distribution that has expectation μ and support with upper bound \bar{v} . First, via similar reasoning, it is without loss to associate each signal with an action recommendation. Second, if a buyer buys, their value should be as high as possible—in this case, equal to \bar{v} (which contrasts with to the setting with information). When the price is p , a buyer that does not buy should have expected value exactly equal to p . Assuming $p < \mu$ (otherwise, the seller would obtain 0 profits via null information), we have the worst case purchasing probability q therefore satisfies $q\bar{v} + (1 - q)p = \mu$. Hence $q = \frac{\mu - p}{\bar{v} - p}$, and profit is $p\frac{\mu - p}{\bar{v} - p}$. The first order condition for p gives:

$$(\bar{v} - p)(\mu - 2p) + p(\mu - p) = 0 \Rightarrow p = \bar{v} - \sqrt{\bar{v}(\bar{v} - \mu)},$$

which yields profit $2\bar{v} - \mu - 2\sqrt{\bar{v}(\bar{v} - \mu)}$. Setting $\bar{v} = 1$ and $\mu = 1/2$ yields the results claimed in the main text. ■

Proof of Claim 3. Suppose that $\mathbb{P}[p_2 < \infty] > 0$. Let the first period distribution be the same as the one that arises in the construction of Proposition 2 (recalling that this allows for random strategies by the seller), which involves $v_1 \in \{w_1, \tilde{w}\}$, which arises from a partitional information structure where w_1 is indifferent between purchasing and waiting with no further information. Consider the following choice of nature for a distribution D_{x,v_1} over values for v_2 :

- If $p_2 > \tilde{w}$, then $v_2 = v_1$.
- If $p_2 \leq \tilde{w}$, then $v_2 = x^2$ with probability $\frac{v_1+x}{x+x^2}$ and $v_2 = -x$ otherwise.

Note that the buyer that arrives in the first period has a value that follows a martingale. For the second buyer, reveal no information if $p_2 \geq \tilde{w}$, and otherwise utilize the worst case partitional information structure given p_2 .

The claim will follow by showing that whenever $\mathbb{P}[p_2 < \tilde{w}] > 0$, the seller's profit from the first buyer can be made arbitrarily small, since the profit from the second buyer is at most $\delta\Pi^*$ (noting that when $p_2 \geq \tilde{w}$ with probability 1, the best the seller can do is Π^* under the constructed value process). Indeed, a buyer with value \tilde{w} has expected value $\tilde{w} - p_1$ from immediate purchase, but value $\delta \cdot \frac{\tilde{w}+x}{x+x^2}(x^2 - \mathbb{E}[p_2 | p_1])\mathbb{P}[p_2 < \tilde{w}]$, which approaches ∞ as x approaches ∞ . In contrast, since the seller charges positive prices, the probability of sale is $\frac{v_1+x}{x+x^2}$, which approaches 0 as x approaches ∞ . The same holds for the buyer who has expected value w_1 . Hence nature can make the probability of sale to any first period buyer arbitrarily small, thereby proving the claim. ■

D. MISCELLANEOUS RESULTS

D.1. Alternative Interpretation

In this section, we consider an information acquisition game for which our solution is of interest. The motivation borrows heavily from Roesler and Szentes (2017). They consider a game with the following timing:

- The buyer first chooses an information structure $\mathcal{I} : V \rightarrow \Delta(S)$.
- The seller then chooses a price $p \in \mathbb{R}$.
- The buyer finally decides whether or not to purchase the object.

It turns out that the resulting information structure results in a payoff for the seller that is optimal given that the information structure is the worst possible, *and assuming that the information structure does not depend on the price.*

For our setting, first consider the $T = 1$ case, and modify the Roesler-Szentes (2017) game so that the buyer's information structure can depend on the price. That is, we take the same timing as above, but allow for information structures of the form $\mathcal{I}(p) : V \rightarrow \Delta(S)$. While in practice it may be difficult to assume that the seller literally chooses this information structure, the information may indeed be provided by a third party (such as Amazon) who would have this power and potentially this objective as well.

Recall Π^* is the seller's maxmin payoff.

Proposition 8. *Consider a one-period setting where the buyer (or third party who acts to maximize the buyer's payoff) chooses the information structure. The seller's profit in the equilibrium of this game is equal to Π^* .*

Note that this does not say the equilibrium is payoff equivalent for the *buyer*. In general it will not be. Still, if one were interested in buyer payoffs, the proposition shows that our analysis is relevant for this case as well. Furthermore, the proof is relatively straightforward.

Proof of Proposition 8. We demonstrate that the optimal choice of information for the buyer results in payoff of Π^* for the seller.

Denote by \mathcal{I}_\emptyset the completely uninformative information structure, and let $\mathcal{I}^*(p)$ be the worst case information structure for the seller when the price chosen is p . The buyer chooses the following information structure:

$$\mathcal{I}(p) = \begin{cases} \mathcal{I}_\emptyset, & p = \Pi^* \\ \mathcal{I}^*(p) & p \neq \Pi^* \end{cases}.$$

The seller chooses price Π^* .

First we compute the profits of the buyer and the seller. Since $\Pi^* < \mathbb{E}_{v \sim F}[v]$ whenever F is non-degenerate, trade happens with probability 1. Since trade is always efficient, total surplus is $\mathbb{E}_{v \sim F}[v]$. We thus have the buyer's surplus is $\mathbb{E}_{v \sim F}[v] - \Pi^*$ and the seller's surplus is Π^* .

Suppose there were a choice of information structure for the buyer which obtained some payoff $u > \mathbb{E}_{v \sim F}[v] - \Pi^*$. Note that the seller's payoff must be at least Π^* in any equilibrium, since this is defined to be the maxmin profit. Hence we have that total surplus in this mechanism is larger than $\mathbb{E}_{v \sim F}[v]$, which is a contradiction. Therefore, the conjectured information structure is optimal.

But if the buyer obtains $\mathbb{E}_{v \sim F}[v] - \Pi^*$, the seller must obtain exactly Π^* ; again, if the seller obtained more, total surplus would be larger than $\mathbb{E}_{v \sim F}[v]$, a contradiction. Hence the seller's payoff in this game is Π^* . ■

Note that, while the buyer obtains different payoffs than in the maxmin benchmark we previously analyzed, they can still be computed simply once we have found Π^* ; since trade always occurs, the buyer obtains $\mathbb{E}_{v \sim F}[v] - \Pi^*$. Note that the same argument also goes through when there are production costs, provided that trade is ex-ante efficient.⁴⁵

The same construction works for an arbitrary horizon, replacing the choice of a single price with the choice of a constant price path of Π^* . The proof is identical, considering the discounted buyer surplus instead of an individual buyer's surplus.

D.2. Buyer Uncertainty

One may wonder why the sellers in our model are so much better informed than the buyers, particularly over information in the far future. While this concern may also apply to much of the literature on robustness in mechanism design, we admit that the reader may find it particularly salient here since buyer knowledge of all future information buyers is particularly demanding.

Part of the difficulty in allowing for buyer learning in contrast to seller learning is that non-Bayesian updating is significantly more complicated (which is also why the commitment benchmark is a more natural starting point than non-commitment). Nevertheless, we seek to accommodate this as follows. Recall that a *dynamic information structure* \mathcal{I} is a sequence of signal sets S_t and probability distributions $I_{a,t} : R_+ \times S_a^{t-1} \times P^t \rightarrow \Delta(S_t)$, for $1 \leq t \leq T$. We now assume that nature has the ability to choose a set Ω , which we take to be a finite set, and dynamic information structures $\{\mathcal{I}^\omega\}_{\omega \in \Omega}$, with associated probability distributions $I_t^\omega : R_+ \times S^{t-1} \times P^t \rightarrow \Delta(S_t)$. Crucially, note that the signal set *does not* depend on ω , as the interpretation is that the buyer observes signal s_t in period t . Let us denote by $\Omega(s^t)$ the set of information structures the buyer believes is feasible after signal history s_t . This allows us to describe which information structures the buyer can “rule out” over time.

For simplicity, we assume that the buyer utilizes a maxmin objective, though similar arguments would apply for general uncertainty averse preferences. It is easy to show that without any restrictions, allowing for arbitrary uncertainty aversion can eliminate any seller profits.

Proposition 9. *Suppose nature can choose any set of dynamic information structures. Then the seller's worst-case single period profit (as well as dynamic profit) is equal to 0.*

⁴⁵The same equivalence fails for Roesler and Szentos (2017) in general when there are costs.

Proof. Consider any price path $(p_t)_{t=1}^T$, and let $\underline{p} = \min_t p_t$. Suppose $\underline{p} > 0$. Let nature choose $S_1 = \{0, 1\}$, $\Omega = \{0, 1\}$.

- At time 1, if $\omega = 1$, the buyer observes $s_1 = 1$ if $v > \underline{p}$ and $s_1 = 0$ otherwise.
- At time 1, if $\omega = 0$, the buyer observes $s_1 = 0$ if $v > \underline{p}$ and $s_1 = 1$ otherwise.

In this case, since the buyer is maxmin, the expected payoff from purchasing is negative, hence they will never purchase. If $\underline{p} = 0$, define $\tilde{p} = \min_{p>0} p_t$. Repeating the same argument above, replacing \underline{p} with \tilde{p} , completes the proof. ■

This result is straightforward, particularly since the martingale condition of expectations is what allowed us to avoid degenerate solutions in the first place. Hence some restriction must be made an updating, and the following appears to be the most-reasonable.

Definition 2. *The buyer is a **within-period Bayesian** if all information structures $I_t^\omega(v, s^{t-1}, p^t)$ with $\omega \in \Omega(s^{t-1})$ are identical.*

In words, a within-period Bayesian does not face any uncertainty over the information they have obtained up until that time. However, they may be non-Bayesian over information they receive in the future. This rules out creating arbitrarily pessimistic Hence information still has the same bite in terms of avoiding degeneracy, but no longer imposing that the buyer has significantly more knowledge of future information than the seller.

Under the within-period Bayesian assumption, Proposition 2 carries through identically, since this proposition only requires us to construct an information structure that holds the seller down to the one-period benchmark. We can also show the following:

Proposition 10. *Under the assumption that all buyers are within period Bayesians, the optimal seller strategy is a constant price path of p^* , delivering discounted profits of $\Pi^* \frac{1-\delta^T}{1-\delta}$.*

Proof of Proposition 10. We consider the same replacement as in Proposition 2, except we assume in addition that nature reveals the true information structure to a buyer at time 1 (in addition to pushing signals forward). Obedience and algebra still demonstrate that:

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

Since $\mathbb{E}[v \mid s_1] = \mathbb{E}^{s_2, \dots, s_T} [\mathbb{E}[v \mid s_1, s_2, \dots, s_\tau]]$ in the case where there is no uncertainty over future information, we have $\mathbb{E}[v \mid s_1] \geq \mathbb{E}^{s_2, \dots, s_T} [\mathbb{E}[v \mid s_1, s_2, \dots, s_\tau]]$. Hence we still obtain the crucial inequality:

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

Since the buyer is a Bayesian in the transformed setting, the same analysis applies and we have $p_1 \geq \mathbb{E}[v \mid s_1, \underline{s}]$. Hence the replacement is obedient, so the probability sale is $\mathbb{E}[\delta^{t-1}]$, according to the realized information structure (noting that this is the parameter that matters for the seller's profit). Hence the seller is made no better off with the replacement. ■

We make two observations on general features of this model, where we do not restrict buyer uncertainty to being of the maxmin variety: First, our result that the worst-case for the seller is a Bayesian buyer under within-period Bayesianism should hold under more general models of ambiguity aversion than the one we present here. Indeed, Riedel (2009) derives a version of the Optional Sampling Theorem for multiple-prior supermartingales, in dynamic models of ambiguity where the agent is time-consistent.

Second, we should *not* expect our results to hold in cases where the buyer is ambiguity *loving* but a within-period Bayesian. In that case, even in the case where the buyer is a within-period Bayesian, nature can induce delay by utilizing ambiguity over the buyer's future information. In that case, preventing the buyer from purchasing in the future by raising the price may help profits, even potentially at the expense of excluding future buyers, in contrast to our main results.