Repeated Sales with Multiple Strategic Buyers

Nicole Immorlica[*] Brendan Lucier[†] Emmanouil Pountourakis[‡]
Samuel Taggart[§]

January 18, 2020

Abstract

In a market with repeated sales of a single item to a single buyer, prior work has established the existence of a zero revenue perfect Bayesian equilibrium in the absence of a commitment device for the seller. This counter-intuitive outcome is the result of strategic purchasing decisions, where the buyer worries that the seller will update future prices in response to past purchasing behavior. We first show that in fact almost any revenue can be achieved in equilibrium, but the zero revenue equilibrium uniquely survives natural refinements. This establishes that single buyer markets without commitment are subject to market failure. However, our main result shows that this market failure depends crucially on the assumption of a single buyer. If there are multiple buyers, the seller can approximate the revenue that is possible with commitment. We construct an intuitive equilibrium for multiple buyers that survives our refinements, in which the seller learns from past purchasing behavior and obtains a constant factor of the per-round Myerson optimal revenue. The seller's pricing policy has a natural explore-exploit structure, where the seller starts with low prices that gradually ascend to learn buyers' values, and in later rounds exploits the surviving high-valued buyers. The result resembles an ascending-price auction with a supply limit, implemented over time.

On face, our result runs counter to intuition from the Coase conjecture in the durable goods literature [Coase 1972] which states that in the absence of commitment, one should expect the VCG outcome (which, for digital goods, would yield trivial revenue for the seller). We argue that our positive result is driven by the assumption of anonymous prices. We show that if the seller is permitted to offer different prices to each agent then the Coasian intuition from the single-item setting binds once more: the seller is no longer able to extract nontrivial revenue from any equilibrium with sufficiently natural structure. In other words, the restriction of the seller to an anonymous price was crucial in deriving nontrivial revenue with unlimited supply. Intuitively, an anonymous price mitigates the ability of the seller to use the information an individual buyer leaks with each purchasing decision. Consequently, buyers are more willing to make nontrivial purchasing decisions, which in turn allows the seller to learn.

1 Introduction

It is now commonplace for regular, repeated purchases to be made through large online platforms. New parents purchase diapers monthly through Amazon Prime. Firms buy online advertising

[*]Microsoft Research
[†]Microsoft Research
[‡]Drexel University
[§]Oberlin College

space millions of times per day through Google, Microsoft and other advertising markets. City-dwellers use delivery services like Foodler and Instacart to purchase their meals and groceries. Each platform is a cornucopia of data, since they can readily observe how pricing decisions affect the purchasing behavior of customers, both in aggregate and individually. It is tempting for a platform to exploit this historical data, by using the past behavior of individual users to tune prices and maximize revenue. However, using revealed preference data in this way runs afoul of game-theoretic considerations. If a regular customer knows that their behavior will impact the prices they will be offered in the future, they will naturally respond by changing their behavior. It is therefore crucial to understand how forward-looking customers will respond to price-learning algorithms, and the implications for how a seller should use historical data to make pricing decisions.

Consider the following simple and fundamental instantiation of the repeated-sales problem, coined the "fishmonger problem" in Devanur et al. (2015). There is a single seller, and each day the seller has a single copy of a good to sell. There is a single buyer, who has a private value $v \geq 0$ for obtaining the good each day, drawn from a distribution known to the seller. Crucially, the value does not change from one day to the next; the buyer has the same value for consuming the good on every day. Each day, the seller posts a take-it-or-leave-it price, and the buyer can choose to accept or reject. The seller is free to set each day's price however she chooses, given the past purchasing behavior of the buyer. On any day that the buyer rejects, the good expires and the seller must discard it. The game is played for infinitely many rounds; the buyer wishes to maximize total time-discounted utility, and the seller wishes to maximize total time-discounted revenue.

How should the seller set her price? If there is only a single round, the well-known solution is to post the Myerson price for the buyer's distribution, which maximizes expected revenue. In the dynamic setting, however, we cannot expect the seller to post the Myseron price each round. After all, if the buyer chose not to purchase on the first day, the seller would naturally want to learn from this information and set a lower price on the following day. It is tempting to guess that the seller can benefit from this opportunity to learn, by offering a variety of prices to gain information about the value v, then use this knowledge to set an aggressive price just below v. However, a surprising folklore result implies that such techniques can never be beneficial to the seller: the average per-round revenue can never be higher than the one-round Myerson revenue. Intuitively, the issue is that a rational buyer would respond to an explore/exploit strategy by pretending at first to have a low value, passing up some opportunities to buy the item, in order to secure a lower price later on. Indeed, this strategic demand-reduction behavior is the essence of bargaining, and is commonly observed in practice.

So what *can* the seller do? To disentangle the strategic behavior of the buyer and seller, it is necessary to study equilibria. Since ours is a repeated game with private information, the appropriate solution concept is perfect Bayesian equilibrium (PBE). A formal definition is given in Section 2, but roughly speaking a PBE requires that the decision taken by each player at each point in time, for any observed history of prices and purchases, is a best response to the anticipated future behavior of the other player, given the seller's belief about the private value (which will depend on the observed behavior of the buyer). Determining how the seller should set prices then reduces to understanding the structure of PBE. Sadly, prior work on equilibria for repeated sales have mostly generated negative results. In particular, there exist PBE in which the seller posts a price equal to the minimum value in the support of the buyer's distribution in every round Devanur et al. (2015); Hart and Tirole (1988); Schmidt (1993). For example, if the buyer's value is supported on $[0,1]$, then there is a PBE with zero revenue for the seller. This extreme and counterintuitive equilibrium is driven by a self-fulfilling prophecy: the buyer never accepts any positive price out of fear that doing so will lead the seller to charge very high prices in the future; as a result, the seller infers that only a buyer with very high type would ever accept a positive price, so the seller would indeed charge very high prices in response. The formal details of the equilibrium are described in Section 3. This construction illustrates that in the absence of commitment power, a seller might suffer extremely low revenue in long-term interaction with a buyer. We note that this conclusion is reminiscient of the Coase conjecture; the primary difference is that the Coase conjecture refers to a durable good that a buyer will purchase only once, whereas in the fishmonger problem the good is perishable and can be repurchased each day (Coase, 1972).

This result is quite negative, but also unsatisfying since the low-revenue equilibrium does not appear to be predictive of real-world outcomes. Why don't we see this behavior in practice? One simplifying assumption in the model is the presence of only a single buyer. Indeed, because there is only one buyer, it is possible for the seller to exploit the buyer's revealed preference in a very targetted way. In contrast, if the seller continues to sell by posting a single price, but that price will be faced by *multiple* buyers, then the opportunity for price-discrimination is diminished. Intuitively, in a market with multiple buyers, each buyer is less worried about being exploited directly, and competition gives an extra incentive to purchase even though this is revealing a signal to the seller. We therefore ask: would the presence of multiple buyers change the structure of equilibrium?

1.1 Our Results

Understanding One Buyer The existence of a zero-revenue equilibrium is discouraging, but we begin by showing that the single-buyer situation is even more dire than that. One might wonder whether the low-revenue equilibrium is simply an edge case, and that better and more plausible equilibria exist. Indeed, we establish a folk theorem that implies that any amount of revenue between the trivial lower bound (that of posting the minimum-supported value every round) and that of Myerson pricing every round can be realized at a PBE of the game. However, despite the rich space of equilibria, we prove that the zero-revenue equilibrium is the unique equilibrium that survives a natural refinement of the set of PBE. Specifically, it is the unique equilibrium in which the buyer uses threshold strategies (i.e., on each round and for any offered price, a buyer purchases if and only if their value is sufficiently high), strategies are Markovian on-path (meaning that on the equilibrium path, the players' strategies depend only on their beliefs and the current price, and not the full history of past play), and the seller offers prices in the support of buyers' distributions. The Markovian and threshold refinements have been studied previously in the context of repeated sales (see Fudenberg and Tirole (1983) and Hart and Tirole (1988)), and are natural conditions for "simple" strategies. We interpret this as strong evidence that the zero-revenue equilibrium, and the market failure it implies, is actually a plausible and natural outcome of the single-buyer repeated game.

Multiple Buyers and Digital Goods We next turn to studying a multi-buyer variant of the Fishmonger problem. Suppose now that there are $n \geq 2$ buyers, each buyer's value is drawn iid from a known distribution, and these values are again fixed over all rounds. The seller has one copy per buyer per day, but must post a single, anonymous price each day. Each buyer independently chooses whether or not to purchase each day.[1]

In contrast to the single-buyer variant, we show that the seller can achieve a constant fraction of the benchmark optimal revenue in a PBE that employs threshold strategies and is Markovian on-path, surviving our refinements. The equilibrium we construct has a natural form, based upon an explore-exploit structure. The seller starts by setting a low price, and slowly raises the price over time. Once sufficiently many buyers stop purchasing, the seller switches to an exploitation phase in which she posts the highest price at which she believes the remaining agents are guaranteed to buy. Since agents are guaranteed to buy, the seller stops learning information about the buyers' values, and posts the same price every round from that point onward.

This equilibrium structure sets up a natural optimization problem for the seller: how quickly should prices be increased, given the way that rational buyers will respond at equilibrium? Typical of explore/exploit algorithms, the seller must balance the rate of learning with the revenue ultimately generated in the exploitation phase. We derive a candidate policy that lowerbounds the seller's revenue, providing an approximation result: if the distribution over buyers' values is regular and has a monopoly price which is bounded away from the extremes of the distribution, then we show how to compute prices (and the corresponding equilibrium thresholds for the buyers) that generate a constant fraction of the Myerson optimal revenue.

Coasian Intuitions and Anonymous Pricing At first glance, the positive results for the multi-buyer game are surprising. They imply that the apparent failure of the single-buyer game for the seller was a fluke, and not robust to a change in the model. Moreover, intuition from the literature on durable goods (Coase, 1972) hints at a different outcome. Coase (1972) suggests that the intuitive prediction for a setting where the seller cannot commit to a mechanism should be that equilibrium implements the efficient VCG outcome. In the single-buyer setting, this produces the trivial equilibrium we show to be focal. For the setting with many buyers, the unlimited nature of the supply should imply a similar conclusion. The explore-exploit structure of our multibuyer equilibrium, however, resembles an ascending-price auction, implemented over many rounds of play, implementing an outcome with greater competition and inflated revenue.

This suggests a larger question: when is it correct to predict the efficient outcome in repeated sales? We consider the role in the above of the exogenous constraint that the seller post an *anonymous* price. What if the seller is able instead to post a separate price for each buyer in a digital goods setting? By slightly strengthening our refinements, we show that the focal equilibrium is the one in which the selling problem decouples for each buyer, and the seller's revenue is again the trivial revenue of the efficient outcome. This is evidence that the anonymous pricing constraint was crucial for the seller's ability to obtain nontrivial revenue in the digital goods setting. Intuitively, the anonymous price made information leakage from nontrivial purchasing strategies less costly to the buyers, which in turn allowed the seller to learn in equilibrium.

Limited Supply We finally consider a two-buyer version of our model where supply is exogenously limited to one item. If multiple buyers wish to purchase at the offered price on a given day, then one of the accepting buyers is chosen uniformly at random to make the purchase. We again exhibit an equilibrium which survives our refinements and in which the seller obtains a constant fraction of the optimal revenue, assuming distributions satisfy the monotone hazard

[1]We choose to model the fishmonger problem as a pricing problem, as this is a common approach taken in practice. We note that one could alternatively model it as a general mechanism design problem, which we leave as a direction for future research.

rate condition. Our equilibrium shares a similar structure with that of the digital goods setting: the seller gradually increases the price until one or both of the buyers drops out, then exploits the buyers for the rest of the game.

1.2 Related Work

Hart and Tirole Hart and Tirole (1988) initiated the study of repeated sales ("rental," in their terms) with a single buyer and a large but finite horizon. They consider a special case with just two possible values. They show that in equilibrium the seller will always post the smaller value for all but a constant number of final rounds. Schmidt (1993) genereralized their result to general discrete distributions. For a survey of this work and the large body of work on closely related models, see the survey of Fudenberg and Villas-Boas (2006). Some variants include Kennan (2001) and Battaglini (2005) who analyze the setting where the value of the buyer is not constant but evolves according to a Markov process, and Conitzer et al. (2012) who study the case where the buyers are short-lived and given the option to anonymize at a cost.

Closest to our work is Devanur et al. (2015), which was the first attempt by the CS community to attempt to move beyond the strong negative results in the setting of Hart and Tirole (1988) and Schmidt (1993), and the first to consider continuous distributions. Like us, they analyzed threshold equilibria, proving that no such equilibria exist for large but finite numbers of rounds. They go on to study the case of partial commtiment, where the seller can commit to never increase the price in the future. They prove existence of PBE for power law distributions and provide revenue guarantees for the uniform distribution $U[0, 1]$. Note that our results can be directly compared to Devanur et al. (2015) where instead of relaxing the commitment assumptions we introduce an extra buyers and an anonymous price, and provide revenue guarantees for a much larger family of distributions.

2 Model

Below, we give the formal description of our model for the sale of a single good to multiple buyers. We assume the seller is exogenously constrained to anonymous pricing. That is, they post a single price each round, which is seen by all buyers. We will consider a seller with many identical items for sale and relax the constraint of anonymous pricing later on, and will describe the formal changes to the model when appropriate.

Game Description and Timing: The dynamic pricing game takes place over an infinite stream of rounds, with time discounting. We consider a digital goods environment: each round, there is one item copy for sale per buyer, and items must be allocated using a common price among n buyers. Before the game begins, each buyer i draws their value v_i for the good independently and identically from some continuous distribution F with bounded support. We assume F is common knowledge. The value for allocation remains unchanged from round to round. Each round k then proceeds in the following way:

1. The seller chooses a price $p_k \geq 0$, which is posted to the buyers.

2. Buyers simultaneously decide whether to accept p_k.

3. Each agent who accepts is awarded a copy of the item and charged p_k.

Utilities: Agents are risk-neutral expected utility maximizers. Utilities are linear in money, additive across rounds, and discounted by a common discount factor $\delta \in (0,1)$ over time. Formally:

- *Seller:* The seller's utility for an outcome to the above game is $\sum_k \delta^k p_k q_k$, where q_k is the number of items sold in round k.
- *Buyer:* The buyer's utility for an outcome is $\sum_{k\,:\,i \text{ accepts}} \delta^k (v_i - p_k)$.

Note that all of our results, except the revenue analyses of Sections 5.2 and Sections 7.2, hold without modification if the seller holds a different discount factor from the buyers. Moreover, the revenue analyses extend in a natural way.

Information: We assume all information and outcomes are common knowledge, with the exception of buyers' values, which are privately held and unknown by all other agents.

Histories: A *history of play* at round k, denoted h^k is different for buyers and the seller, but generally consists of all past pricing and purchasing decisions. Formally, h^k consists of consists of the vector $\mathbf{p}[k-1] = (p_1, \ldots, p_{k-1})$ of past prices, as well as the purchasing decisions of agents in each past round, denoted $\mathbf{D}[k-1] = (\mathbf{D}^1, \ldots, \mathbf{D}^{k-1})$, where $\mathbf{D}^j = (D_1^j, \ldots, D_n^j) \in \{A, R\}^n$ is the vector of accept/reject decisions for each agent i in round j.

Beliefs: The seller does not know any buyer's values, and buyers only know their own. As mentioned earlier, this uncertainty is modeled with a Bayesian prior. After every round of play, the actions of agents may reveal information about their private values, and hence agents' beliefs must be updated. We consider only outcomes where agents' posteriors after each round are shared, which is possible because all actions are commonly observed. The prior for v_i after history h^k, denoted $\mu_i^k(\cdot \mid h^k)$, is a probability measure over the support of F. The joint posterior at round k is denoted $\boldsymbol{\mu}^k = \times_i \mu_i^k$. After round k, the seller believes values are distributed according to $\boldsymbol{\mu}^k$, and buyer i believes other buyers' values are distributed according to $\boldsymbol{\mu}_{-i}^k$.

Strategies: Generally, strategies are maps from histories and private information to actions in round k:

- A seller strategy $\sigma_S^k(h^k)$ specifies for every history h^k a nonnegative price p_k.
- Buyer i's strategy $\sigma_i^k(h^k, p_k; v_i)$ specifies for every buyer history a response to price p_k for every possible value of buyer i.

Equilibrium: Our solution concept is Perfect Bayesian Equilibrium (PBE). Perfect Bayesian Equilibrium imposes joint requirements on beliefs and strategies: beliefs must be updated accurately given strategies, and given beliefs, strategies must be sequentially optimal for all agents. Formally, a profile of strategies $\boldsymbol{\sigma} = (\sigma_S^k(\cdot), \sigma_1^k(\cdot), \ldots, \sigma_n^k(\cdot))$ and beliefs $\boldsymbol{\mu}^k(\cdot)$ for $k \geq 0$ is a PBE only if:

- *Bayesian updating:* For every history h^k, if there is some v such that $\mu_i^k(v \mid h^k) > 0$ and $\sigma_i^k(h^k, p_k; v) = D_i^k$, then $\mu_i^k(v \mid h^k)$ is computed according to Bayes' rule. Importantly, for histories which would not occur according to $(\sigma_S^k(\cdot), \sigma_1^k(\cdot), \ldots, \sigma_n^k(\cdot))$ under any realization of buyers' values, beliefs may be arbitrary.

- *Sequential optimality:* Let $u_S(\boldsymbol{\sigma} \mid h^k, \boldsymbol{\mu}^k)$ denote the expected utility of the seller from the continuation of the game from stage k according to $\boldsymbol{\sigma}$, given that buyers' values are distributed according to $\boldsymbol{\mu}^k(h^k)$. We require that for every alternate strategy σ'_S of the seller, we have that $u_S(\boldsymbol{\sigma} \mid h^k, \boldsymbol{\mu}^k) \geq u_S(\sigma'_S, \boldsymbol{\sigma}_{-S} \mid h^k, \boldsymbol{\mu}^k)$. Similarly if $u_i(\boldsymbol{\sigma} \mid h^k, \boldsymbol{\mu}^k, p_k; v_i)$ is the expected utility of a buyer with value v_i offered price p_k under history h^k and beliefs $\boldsymbol{\mu}^k_{-i}(h^k)$ on other buyers' values, $u_i(\boldsymbol{\sigma} \mid h^k, \boldsymbol{\mu}^k, p_k; v_i) \geq u_i(\sigma'_i, \boldsymbol{\sigma}_{-i} \mid h^k, \boldsymbol{\mu}^k; v_i)$ for every alternate strategy σ'_i.

Simple Equilibria: Equilibria may in general be extremely complicated. We focus on equilibria satisfying two refinements:

- *Markovian on path:* An equilibrium is *Markovian on path* if on the equilibrium path, agents condition their actions only on the public beliefs and their private information, rather than the complete history. Formally, for any profile of buyer values $\mathbf{v}$ and strategy profile $\boldsymbol{\sigma}$, let h^k and $h^{k'}$ be the histories generated by $\boldsymbol{\sigma}$ under $\mathbf{v}$. If $\boldsymbol{\mu}^k = \boldsymbol{\mu}^{k'}$, then $p_k = p_{k'}$ and $\mathbf{D}^k = \mathbf{D}^{k'}$.

- *Threshold equilibrium:* If a buyer will buy when their value is v_i, they will also buy with any higher value. Formally, a PBE is a *threshold equilibrium* if for each history h^k and price p_k, there is a threshold $t_i(h^k, p_k)$ such that for each agent i, i accepts p_k if and only if $v_i \geq t_i(h^k, p_k)$. Note that in threshold equilibria, updated beliefs derived from on-path histories will be the value distribution F conditioned to some interval $[a, b]$ for each agent. For such equilibria, we will therefore summarize beliefs over agent i's value with the notation F^b_a to denote F conditioned to the interval $[a, b]$.

We refer to threshold equilibria which are Markovian on path as *simple*. Note that simplicity is a refinement rather than a restriction of the strategy space.

3 Folk Theorem

We first explore the space of Markovian on path threshold equilibria with no further refinements. It is well-known from previous work on the subject that there exists an equilibrium for the one-buyer case in which the seller gets no revenue and does not learn anything about the buyer's value. The buyers refuse all positive prices, and deviation is punished by the seller with high prices in the future. We refer to this as the no-learning equilibrium, and for completeness present the equilibrium in Appendix A. Formally, we have:

Theorem 1 (Devanur et al. (2015))**.** *For $\delta \geq 1/2$ and any number of buyers there is a simple PBE in which the seller does not learn, and posts a price of* 0 *every round. All buyers accept each round.*

The no-learning equilibrium is considered unnatural and unpredictive. In this and the next section, we offer a more nuanced view. We prove a folk theorem: the no-learning equilibrium can be used to enforce other even less intuitive equilibria, including posting any fixed price every round. In other words, PBE is ineffective at ruling out commitment. We solve this problem in Section 4, by offering an additional, intuitive refinement which surprisingly eliminates all equilibria but precisely the no-learning equilibrium. This suggests that such behavior is a reasonable outcome to the game.

Theorem 2 (Folk theorem)**.** *If $\delta \geq 1/2$, then for any price p, there is a Markovian on path threshold PBE of the dynamic pricing game with n buyers where the seller offers price p every round on the equilibrium path, regardless of the action of the buyers. This holds regardless of the initial prior over buyers' values.*

We prove the theorem in Appendix B. Intuitively, we use the no-learning equilibrium to commit the seller to a strategy. One way to understand the space of PBE is in terms of pairs of attainable payoffs for the buyers and the seller. Theorem 2 implies that the Pareto frontier of attainable payoffs under our two simplicity refinements is at least as strong as that attainable from posting the same price each round. A natural question is whether there are PBE which surpass this frontier. The best known bounds on the performance of PBE is a theorem due to Devanur et al. (2015), which we rephrase below.

Theorem 3 (Devanur et al. (2015))**.** *For any target total expected buyer utility U and revenue R attainable in a PBE, there is a mechanism for the single-round game in which the buyers attain total expected utility $(1-\delta)U$ and the seller attains expected revenue $(1-\delta)R$.*

The proof is constructive: given the PBE attaining R and U, the mechanism designer may in essence simulate the PBE on the reported values of the buyers. In other words, PBE resemble single-shot mechanisms with stronger incentive constraints. Theorem 3 implies that the utility-revenue Pareto frontier for PBE cannot generally exceed that of the single-shot mechanism design problem. For one buyer, Theorem 3 implies that the folk theorem is tight - the utility and revenue guarantees are the best possible.

Theorem 2 implies a troubling multiplicity of equilibria, all with very different outcomes for both the seller and the buyers. It implies that further study of PBE is not worthwhile without a manner of refining away equilibria. We provide such a selection tool in the next section.

4 Non-Robustness of One-Buyer Learning Equilibria

We now specifically consider the case of one buyer and one seller. In this setting, Theorem 2 proves that there are Markovian on path threshold equilibria which are totally efficient, totally inefficient, and revenue-maximizing, as well as everything in between. We argue that all inefficient equilibria exhibit unnatural seller behavior. In particular, in any equilibrium of Theorem 2 where the seller posts a non-trivial price, there are continuations in which the seller offers prices which will be accepted with probability zero according to the current beliefs. We prove in this section that every simple equilibrium of the one-buyer case, except those in which all buyers accept every round and no learning occurs on the part of the seller requires such odd behavior. This leaves only equilibria in which the seller posts a price at or below the bottom of buyers' common support each round. We first formalize "natural" seller behavior.

Definition 4. *A perfect Bayesian threshold equilbrium $\boldsymbol{\sigma}$ of the single-buyer has* natural prices on-path *(or simply* natural prices*) if for every on-path history h^k with beliefs over buyer values supported on $[a_i, b_i]$ for $i \in \{1, \ldots, n\}$, the seller's price $\sigma_S^k(h^k)$ lies in $\cup_i [a_i, b_i)$.*

In other words, the seller does not offer any prices which would not be obviously dominated for any joint distribution over buyer values with the given support. Though this requirement might seem mild, it in fact suffices to eliminate all nontrivial equilibria.

Theorem 5. *In the single-buyer game, let the initial value distribution F be supported on $[a, b]$. If $\delta > 1/2$, then in any simple PBE with natural prices, the seller posts a price which is at most the support lower bound, a, every round, which is accepted by all types. In other words, no learning will occur.*

Driving the proof of Theorem 5 will be an idea from single-dimensional mechanism design: in equilibrium, allocations are monotone in type. In the repeated pricing setting, agents are maximizing their total discounted utility, which is a function of total discounted allocation and total discounted payments. These quantities satisfy the usual incentive constraints from mechanism design, and hence intuitions from mechanism design carry over. We now define these formally:

Definition 6. *Given a PBE of the single-buyer game, let: $x_k(v)$ be an indicator variable of whether or not the buyer with value v purchases in round k on the equilibrium path, and let $p_k(v)$ be the on-path price offered that round. Then we may define the following:*

- *The* total discounted allocation*:*

$$X(v) = \sum_{k=1}^{\infty} \delta^{k-1} x_k(v)$$

- *The* total discounted payments*:*

$$P(v) = \sum_{k=1}^{\infty} \delta^{k-1} p_k(v) x_k(v)$$

- *The* total discounted utility*:*

$$U(v) = vX(v) - P(v)$$

.

Lemma 7. *In any PBE of the one-buyer game, the total discounted allocation, payments, and utility (respectively $X(v)$, $P(v)$, and $U(v)$), are nondecreasing in v.*

To prove this claim, we invoke a theorem of Myerson (1981):

Theorem 8 (Myerson (1981))**.** *Let $f(\cdot)$, $g(\cdot)$ and be functions from some interval $[a, b]$ $(a > 0)$ to the positive reals, and assume the following holds for all v and v' in $[a, b]$:*

$$vf(v) - g(v) \geq vf(v') - g(v').$$

Then the following must be true:

1. *$f(\cdot)$ is nondecreasing in v.*
2. *$g(v) = vf(v) - \int_a^v f(s)\, ds + g(a)$.*

The classic application of this theorem sets $f(\cdot)$ to be the equilibrium allocation probability in a single-item auction and $g(\cdot)$ the equilibrium expected payments. We take a similar approach to prove Lemma 7.

Proof of Lemma 7. Consider an buyer with value v, who must choose a strategy. Among their options are to pretend to have a different value, say v', and play the actions that value would play. Doing so would yield total discounted allocation $X(v')$, total discounted payments $P(v')$, and total discounted utility $U(v')$. Since the buyer is best responding, it must be that $vX(v)-P(v) \geq vX(v') - P(v')$. We may now invoke Theorem 8. Monotonicity of $X(\cdot)$ follows from part (1) of the theorem, and monotonicity of $P(\cdot)$ from part (2). Noting that $U(v) = \int_a^v X(s)\,ds - P(a)$ shows $U(\cdot)$ to be nondecreasing as well. □

We will only use monotonicity of allocations here. In Appendix C, we make heavier use of Lemma 7 to derive alternate sufficient conditions under which the conclusions of Theorem 5 hold.

We now show that natural prices induces non-monotonicity around any threshold t other than the bottom of the support. In particular, we will show that there must exist a type below t with high cumulative allocation, while the threshold type gets allocated strictly less often. This contradicts Lemma 7.

Lemma 9. *For any $\delta > 1/2$, consider any simple PBE of the single-buyer game satisfying natural prices with distribution supported on $[a, b]$ and first-round threshold $t > a$. There exists a type $t' < t$ such that $X(t') > 1$.*

Proof. We argue by contradiction. We will assume that for all $t' < t$, $X(t') \leq 1$ and use natural prices, along with simplicity of equilibrium, to show that there is at least one type less than t who would prefer to deviate from the equilibrium.

We first argue that we may assume the existence of some M such that all types in $[a, t)$ have rejected by round M. Assume this is not the case. Then let k_ϵ be the earliest round such that all agents in $[a, t - \epsilon)$ have rejected at least once. If it is the case that $k_\epsilon \to \infty$ as $\epsilon \to 0$, then because $\delta > 1/2$, it must be the case that there exists some $t' < t$ with $X(t') > 1$, which would prove the lemma. Hence we may assume that the number of rounds before every type in $[a, t)$ would reject at least once on the equilibrium path is finite.

Let M an index such that all agents in $[a, t)$ have rejected before round M. We now claim that there is a round $M^* \leq M$ such that a positive measure of types accept in every one of rounds $1, \ldots, M^* - 1$, but all such agents reject in round M^*. If not, then it must be that a positive measure of agents accept in every round up to and including M, a contradiction. Let the interval of such agents be $[a^*, t)$. (The upper bound being t is implied by the threshold property.)

Finally, we show that the existence of M^*, combined with natural prices and the Markovian on path property, implies a profitable deviation for some buyer with type in $[a^*, t)$. First note that the beliefs conditioned on acceptance in rounds $1, \ldots, M^* - 1$ do not change after round M^*, as all agents who accepted in rounds $1, \ldots, M^* - 1$ will reject in round M^*. Because beliefs don't change, the Markovian in path property implies that actions don't change - hence, in this continuation, no agent in $[a^*, t)$ accepts after round $M^* - 1$. On the other hand, the requirement of natural prices on path implies that the seller offers a price $p \in [a^*, t)$ in every round after $M^* - 1$. Some type in (p, t) would clearly prefer to accept at least once rather than reject forever, yielding a contradiction. □

Proof of Theorem 5. Fix a $\delta > 1/2$, and consider a round of the game in which the beliefs are supported on $[a, b]$ and for which the buyer has a nontrivial threshold t (i.e. above the bottom of the support of the current beliefs). Subgame perfection implies that we may assume this round is the first. We know from Lemma 9 that there is a value $t' < t$ such that $X(t') > 1$. We will show that we may break ties so that $X(t) = 1$, which contradicts Lemma 7.

By the definition of threshold equilibrium, the buyer with type t accepts this round. Natural prices implies that upon seeing an acceptance, the seller will never price below t. It follows that the buyer with value t will not get utility from any subsequent round. We may therefore assume they reject in every round without changing their utility. Moreover, such tiebreaking doesn't change the incentives of the seller, as the type t buyer has measure 0. Hence, there is an equilibrium with $X(t) = 1$ and $X(t') > 1$ for some $t' < t$, contradicting Lemma 7. □

In Appendix C we give an alternate refinement which similarly eliminates learning equilibria. Theorem 5 and Appendix C together strongly suggest that with just one buyer, one should not expect the seller to learn from purchasing behavior. This strengthens the conclusions of Hart and Tirole (1988) and Schmidt (1993) and extends them to continuous distributions. In Section 5.1, we show that these conclusions are critically dependent on the presence of only a single buyer. With multiple buyers, we give a simple PBE with natural prices in which the seller learns from buyers' actions. Moreover, the seller is able to use this knowledge to obtain revenue comparable with the revenue of the optimal auction run in every round.

5 Multi-Buyer Equilibrium

In what follows, we describe a family of equilibria with $n \geq 2$ buyers whose values are independent and identically distributed, with valuation function F and discount factor $\delta \geq 1/2$. These equilibria have two desirable properties. First, they survive the refinements proposed in the previous section, and can therefore be considered robust. Second, and in contrast to the single-buyer case, we show that the seller gets nontrivial discounted revenue. We present the main ideas of the equilibrium in Section 5.1, and leave the formal description to Appendix D. In Section 5.2, we discuss revenue guarantees for our equilibria.

5.1 Supply-Limiting Equilibrium

In this section, we informally describe our family of equilibria. A full version of the equilibrium, with formal descriptions of strategies, can be found in Appendix D. The equilibrium consists of two phases: an *exploration phase*, followed by an *exploitation phase*. In the initial exploration phase, the seller starts the price low, and gradually raises it over time. Each round, a subset of the buyers play nontrivial threshold strategies. Those who reject are priced out for the remainder of the game. Those who accept go on to the next round. Through this process, the seller progressively winnows out low-valued buyers, while a set of targeted buyers S remains. The exploration phase ends when the seller has learned a sufficient amount about the targeted buyers in S. We discuss precise conditions shortly.

Following the exploration phase is an exploitation phase, where the seller ceases to learn, instead exploiting their current knowledge for revenue. At the end of the exploration phase, all buyers in the targeted set S share the same history of accepting every past price (except possibly the final price, in the case that all buyers in S reject). Hence, the seller's posterior over their values is identical, supported on some interval $[a, b]$. The seller now posts single price for the remainder of the game, e.g. a. The incentives of this phase closely resemble the zero-learning equilibrium.

The equilibrium structure described above is compatible with a variety of stopping rules governing the transition between exploration and exploitation. As examples, the seller could cease exploration when the price ascends above a certain price target p^t, or when the number of buyers who accepted in the previous rounds falls below a certain number k. The equilibria we formalize in Appendix D combines these rules, stopping when the price is high *or* the number of remaining buyers is less than k. We refer to k as the *supply limit* of the equilibrium, based on the observation that the exploration process serves to artificially increase competition to prepare for the exploitation phase, and we refer to our equilibria of this form as *supply limiting equilibria*. In Section 5.2, we show that there always exists a choice of supply limit k and target price p^t that guarantee an equilibrium revenue that is a constant fraction of the revenue the seller could obtain from offering the Myerson optimal price every round for the entire game.

We now highlight the key aspects of the buyers' and sellers' optimization problems.

Buyers: In the exploration phase, the seller posts progressively higher prices, such that once a buyer i rejects a price that is accepted by any other buyers, then i rejects all prices for the remainder of the game. If, in a given stage of the exploration phase, buyer i has not yet rejected a price, they choose whether or not to accept the current price p according to a threshold policy. Let F be the current beliefs over the value of buyer i, supported on $[a, b]$. As in the single-item equilibrium, the threshold equation will depend on the number of buyers who have not yet rejected. Assume for simplicity that the number of such buyers is n in the current round. Buyer i's threshold t solves the equation:

$$(t - p) = (t - a)\frac{\delta}{1-\delta}F(t)^{n-1}. \tag{1}$$

The left-hand side represents the utility of a buyer with value t who accepts the current price: they get $t - p$ utility from the current round, but every subsequent price will be at least t. The righthand side corresponds to the utility such an agent would obtain from rejecting: they receive no utility from the current round, but if all other agents reject the current price, then the seller will post a for the rest of the game, and the buyer will attain a total discounted utility of $\frac{\delta}{1-\delta}(t - a)$ from future rounds. This latter event occurs with probability $F(t)^{n-1}$.

In the exploit phase of the equilibrium, the seller targets the set of buyers who have rejected the fewest times, and posts the bottom of the support of their current beliefs. These buyers simply accept each round. In other words, these buyers play the no-learning equilibrium with the seller for the remainder of the game, and their incentives are the same as in that equilibrium.

Seller The seller's pricing problem is governed by a simple Markov decision process, in which they must choose a price for the current round (and implicitly, a threshold for the targeted buyers) which optimally trades off revenue in the current round against revenue gains from learning in future rounds. The seller faces a range of prices for which there exist solutions to equation (1). Let $F(\cdot)$ be the belief distribution for set of buyers who have not yet rejected a price, and for simplicity, assume that there are n such buyers. If we define $R(a, b, j)$ to be the discounted expected revenue the seller can obtain in the j-buyer equilibrium with the same supply limit m and beliefs equal to F conditioned to the interval $[a, b]$, then we may write the seller's recurrence for the revenue from posting a price p as:

$$\begin{aligned} R(a,b,p,n) = \quad & F(t(p))^n \tfrac{\delta}{1-\delta} na + \textstyle\sum_{j=1}^{m} \binom{n}{j} (1 - F(t(p)))^j F(t(p))^{n-j} (pj + \tfrac{\delta}{1-\delta} t(p) j) \\ & + \textstyle\sum_{j=m+1}^{n} \binom{n}{j} (1 - F(t(p)))^j F(t(p))^{n-j} (pj + \delta R(t(p), b, j)) \end{aligned} \tag{2}$$

Each round of the exploration phase, the seller chooses a price from among those for which the threshold equation (1) has a solution to maximize the recursively-defined revenue (2).

In the exploitation phase, as mentioned earlier, the seller posts the bottom of the support of the buyers who have rejected the fewest times. The incentives for this case match those of the no-learning equilibrium.

5.2 Revenue Guarantees

We now briefly describe the revenue guarantee for the supply-limiting equilibria in the digital goods model.

Theorem 10. *Assume the initial value distribution F is regular monopoly quantile q^* satisfying $q^* \geq 1/n$. Then as long as $\delta \geq n/(n+1)$, there exists a supply limit k^* and target price p^t such that the k^*-supply limiting equilibrium described in Section 5.1 is a constant approximation to the revenue obtained by running the revenue-optimal auction for F each round.*

The proof of Theorem 10 can be found in Appendix E. We do not derive an explicit constant for reasons of readability. Note that the (unstated) constant is independent of the number of buyers n, and holds for all $\delta \geq n/(n+1)$.

To prove our revenue guarantee, we derive a feasible solution to the seller's dynamic revenue maximization problem. The optimal solution to their revenue maximization problem, i.e. their equilibrium strategy, will only be better. This candidate pricing strategy learns as aggressively as possible, targeting a price which is close to the monopoly price, and a supply limit which is close to the ex ante demand at the monopoly price in the static pricing problem with n buyers. The proof of Theorem 10 shows that this candidate pricing policy we propose reaches the target price with a constant probability, and quickly enough that the discount factor is lower bounded by a constant.

6 Discriminatory vs. Anonymous Pricing

The intuition from the dynamic sale of a durable good (Coase (1972)) is that the only outcomes implementable in PBE are the efficient ones. The ability of the seller to restrict supply and earn nontrivial revenue in the digital goods setting therefore may come as a surprise. In this section, we present theoretical evidence that the force driving these latter results was our exogenous restriction of the seller to posting *anonymous* prices.

To understand the effect of the anonymous prices restriction, we consider a seller who is able to post a different price for every buyer every round. We show that under a slight strengthening of our refinements, the introduction of discriminatory pricing destroys the potential for high-revenue equilibria that we observed in the settings with anonymous pricing. We show that any equilibrium satisfying our refinements necessarily decouples across buyers into the trivial zero-learning equilibrium for each buyer. This strongly suggests that the anonymous restriction on pricing was necessary for the seller to obtain high revenue. Intuitively, an anonymous price lessens the impact an individual buyer's information leakage has on the price. Consequently, the buyer is more willing to leak information by exhibiting nontrivial purchasing behavior. This allows the seller to learn about buyers' values with ascending prices.

To enable us to formally discuss repeated sales with discriminatory pricing, we first lay out additional notation and note the differences between the model with anonymous pricing. The timing of the repeated sales game remains largely unchanged: values are realized at the beginning of the game. Each round, the seller posts prices and buyers decide to accept or reject. The main difference is that the seller now posts a separate price p_k^i to each buyer i in round k (with the full vector of prices for round k denoted $\mathbf{p}_k$. We assume that buyers are able to observe the prices offered to all other buyers, as well as other buyers' purchasing decisions. Consequently, on-path beliefs are still commonly held among the buyers and the seller (with the exception of each individual buyer's knowledge of their own value). Furthermore, buyers may condition their strategies on the prices offered to other buyers, or on other buyers' past purchasing decisions. The negative result of this section holds even if monitoring among buyers is less than perfect, and possibly holds with weaker assumptions.

We now turn to refinements for this model. The threshold equilibrium refinement extends essentially unchanged. It still requires that buyers map each pricing decision by the seller (which for discriminatory pricing means the full vector of prices $\mathbf{p}_k$) to a threshold, which guides that buyer's purchasing decision. Formally:

Definition 11. *A PBE for discriminatory pricing is a* threshold equilibrium *if for each history h^k and price vector $\mathbf{p}_k$, there is a threshold $t_i(h^k, \mathbf{p}_k)$ such that for each agent i, i accepts p_k^i if and only if $v_i \geq t_i(h^k, \mathbf{p}_k)$.*

In previous sections, we considered equilibria which were Markovian on path: the pricing and purchasing decisions depended only on the current beliefs over the values of the buyers, and not over the past history of play. This had the key consequence that if a seller offered a price that all types for all buyers would uniformly accept (or uniformly reject), the seller would not learn, and would consequently post the same price for the rest of the game. For multibuyer discriminatory pricing, we isolate this property:

Definition 12. *A PBE for discriminatory pricing satisfies the* no backtracking property *if for any two consecutive rounds k and $k+1$ such that the beliefs on a buyer i's value are equal, i.e. $\boldsymbol{\mu}_i^k = \boldsymbol{\mu}_i^{k+1}$, then pricing decisions and purchasing decisions for those two rounds must also be equal. That is, $\mathbf{p}_k = \mathbf{p}_{k+1}$ and $\mathbf{D}^k = \mathbf{D}^{k+1}$.*

With multiple buyers, we note that no backtracking is not strictly weaker or stronger than the Markovian on path refinement. For example, it is possible in Markovian equilibria for buyers to condition their purchasing behavior on other buyers' prices and purchases in such a way that they may be offered the same prices with the same beliefs and respond in two different ways. This behavior violates the no backtracking property.

Finally, we extend the notion of natural prices on path to the multibuyer setting:

Definition 13. *A PBE for discriminatory pricing satisfies* discriminatory natural prices on-path *(or* discriminatory natural prices *for short), if for each buyer i, every on-path history h^k with beliefs for buyer i supported on $[a_i, b_i]$, the seller's price p_k^i for buyer i in round k lies in $[a_i, b_i)$.*

One could imagine a weaker restriction, where all buyers' prices are constrained to $\cup_i [a_i, b_i)$. We motivate the use of the stronger Definition 13 by noting that in the single-shot game, the optimal price vector for the seller would lie in $[a_i, b_i)$ for each agent, regardless of the distribution of buyer i's value over that interval. Definition 13 therefore requires the seller's dynamic actions to be plausible strategies for the single-shot pricing game.

We can now characterize sufficiently simple equilibria, i.e. threshold PBE satisfying no backtracking and discriminatory natural prices on path. We obtain the following theorem:

Theorem 14. *For digital goods with discriminatory pricing, let the initial value distribution F be supported on $[a, b]$. If $\delta > 1/2$, then in any threshold PBE satisfying no backtracking and discriminatory natural prices, the seller posts the support lower bound, a, every round to every buyer, and all buyers accept this price, regardless of type. In other words, no learning will occur.*

To prove Theorem 14, we extend the ideas that drove the proof of the single-buyer result, Theorem 5, and show that monotonicity of each buyer's expected total discounted allocation is incompatible with our refinements unless each agent's the threshold is at the bottom of the support of the value distribution. The proof can be found in Appendix F.

7 Epilogue: Exogenously Limited Supply

In Section 5, we gave a family of equilibria where the seller effectively implements a supply limit via an ascending price. One natural question is whether equilibrium qualitatively changes when the seller's supply limit is given up front as an exogenous constraint. In this section, we make a small step towards characterizing such settings. We consider a setting with two buyers and one item for sale. The seller is again constrained to an anonymous price. In any given round, there is now a possibility that both buyers accept a the current price. When this occurs, the item is given uniformly at random to one of the buyers. Otherwise, the game proceeds as it did with digital goods.

In what follows, we describe a simple equilibrium with two buyers whose values are independent and identically distributed, with distribution function F, and discount factor $\delta \geq 2/3$. This equilibrium shares the two desirable properties of our digital goods equilibrium: first, it survives the refinements proposed in Section 4, and can therefore be considered robust. Second, the seller gets nontrivial discounted revenue, which stands in contrast to the robust no-learning equilibrium of the single-buyer case. We present the main ideas of the equilibrium in Section 7.1 and leave the full formal description to Appendix G. In Section 7.2, we derive revenue guarantees for our equilibrium.

7.1 Equilibrium Description

Our equilibrium again consists of an exploration phase and an exploitation phase. In the exploration phase, which starts in the very first round and lasts until one or more buyers reject, the seller offers prices which will be rejected with positive probability. Once an agent rejects, the equilibrium enters the exploitation phase, which lasts until the end of the game. If a single agent triggered the phase by rejecting, the seller ignores this agent, and posts a price at the bottom of the support of the beliefs for the stronger agent. This price is offered and accepted for the rest of the game. If both agents rejected to trigger exploitation, then the seller posts a price at the bottom of the common support of their beliefs. Below, we informally describe the optimization problems of the seller and the buyers to convey the main ideas of the equilibrium.

Buyers In a given round j of the exploration phase, neither buyer has rejected a price yet. The seller offers a new price, say, p_j, and the buyers, whose beliefs are distributed i.i.d. according to some posterior F_j supported on $[a_j, b_j]$, behave according to a threshold $t_j(p_j)$ solving the equation:

$$\left(\frac{1-F_j(t_j)}{2} + F_j(t_j)\right)(t_j - p_j) = \frac{\delta}{2(1-\delta)}(t_j - a_j)F_j(t_j). \tag{3}$$

The lefthand side represents the utility of a buyer with type t_j who accepts the price p_j, which is $t_j - p_j$ times the probability of winning in the current round, with a continuation utility of zero. The probability of winning in the current round is the probability that the other buyer rejects, $F_j(t_j)$, plus the probability that the other buyer accepts divided by two, as the item is allocated uniformly at random. The righthand side represents the threshold buyer's utility from rejection - if the other buyer accepts, then the seller will only post prices above t_j in subsequent rounds yielding zero continuation utility for the threshold buyer. If both buyers reject, then they share the item for the rest of the game at price a_j. Note that the threshold equation (3) only implies that a buyer with value t_j will be indifferent between accepting and rejecting; it does not immediately imply that buyers will value greater than t_j will prefer to accept, or those with lower values will prefer to reject. We prove that this is indeed also the case in Appendix G.

In the exploitation phase, the seller targets the buyers with the strongest value distribution conditioned on past behavior, and prices at the bottom of their support. The buyer incentives in this phase of the game are similar to those of the no-learning equilibrium of Theorem 1.

Seller In the explore phase, the seller's optimization problem is an algorithmic pricing problem. Each round j, the seller must jointly choose a price p_j and threshold $t_j(p_j)$ jointly satisfying the threshold equation (3), for current beliefs F_j supported on $[a_j, b_j]$. They know from the buyers' strategies that such prices will be met with a threshold response. It therefore suffices for the seller to maximize the following value function:

$$\begin{aligned} R(a_j, b_j, p_j) = \quad & (1 - F_j(t_j(p_j)))^2 (p_j + \delta R(t_j(p_j), b_j)) \\ & + 2F_j(t_j(p_j))(1 - F_j(t_j(p_j)))(p_j + \tfrac{t_j(p_j)\delta}{1-\delta}) + F_j(t_j(p_j))^2 (\tfrac{a_j \delta}{1-\delta}), \end{aligned} \tag{4}$$

where $R(x, y)$ is the optimal continuation revenue from the equilibrium with values distributed according to F conditioned to $[x, y]$ and $t_j(p_j)$ is the threshold corresponding to price p_j. The three terms of this function represent the three possible outcomes to the current round: both buyers accept, exactly one accepts, and both reject.

In our presentation of the equilibrium, we leave the specific price path selected by the seller as the implicit solution to the above optimization, and note that any policy for choosing prices and corresponding thresholds satisfying the threshold equation will support threshold behavior by the buyers. To find a policy arbitrarily close to optimal, one may discretize value space and solve the Markov decision problem associated with the value function (4) by value iteration, though we make no claim as to the computational efficiency of this method. For a computationally constrained buyer, we give in the next section a computationally tractiable threshold-supported pricing policy which obtains provably high revenue, and which serves as a lower bound to the seller's revenue in equilibrium.

7.2 Revenue Guarantees

We now argue that if distributions are well-behaved, the revenue of the equilibrium outlined in the previous section (and discussed in full detail in Appendix G) has high revenue. Specifically, we assume that the *hazard rate* $\frac{f(v)}{1-F(v)}$ of the distribution is increasing in v - a standard assumption in mechanism design. Note that a similar holds for any regular distribution, parametrized by the quantile of the monopoly price. As a benchmark, we use the revenue that the seller would obtain if they used the optimal auction in every round (e.g. Myerson (1981)). By Theorem 3, this benchmark is an upper bound on the seller's revenue in any PBE. Our revenue result is the following:

Theorem 15. *Assume the value distribution F of the two buyers has a monotone hazard rate, and assume $\delta \geq 2/3$. In the equilibrium described in Section 7.1 and Appendix G, the seller obtains revenue which is at least $\frac{1}{3e^2}$ of the revenue of the optimal auction run each round.*

The proof follows the same strategy as Theorem 10. To argue the theorem, we first observe that in the exploration phase of the equilibrium, the seller may offer any price which has a threshold response, and the arguments in the previous section ensure that buyers will be incentivized to adhere to threshold responses. It follows that we may analyze any sequence of prices for the exploration phase, and as long as each price has a threshold response, the resulting revenue will be a lower bound on the actual revenue of the seller.

Rather than compare ourselves directly to the benchmark of the optimal single-shot revenue each round, we will upper bound this benchmark and compare our equilibrium's revenue to this upper bound. The optimal single-shot revenue attainable from n IID buyers is concave in n. It follows that twice the optimal single-shot revenue for one buyer, which can be obtained by posting a price, is an upper bound on our revenue benchmark. Formally, our upper bound is $\frac{2(1-F(p^*))p^*}{1-\delta}$, where p^* is the single-buyer monopoly price, which maximizes $(1-F(p))p$.

To relate our equilibrium revenue to this upper bound, we imagine the seller choosing a sequence of prices which increases the threshold quickly until it reaches p^*, after which the seller voluntarily enters the exploit phase. Assuming both agents have value above p^*, the seller will receive revenue of p^* in perpetuity starting as soon as the threshold reaches this point. By upperbounding the time it takes for this to occur, we can lowerbound the expected revenue from this sequence of prices, and therefore the revenue from the price sequence actually selected by the seller in equilibrium.

First, consider an arbitrary round j of the explore phase, where the current beliefs are over an interval $[a_j, b_j]$ with CDF F_j. We argue that there is always a way for the seller to induce a threshold t which learns "quickly." Formally:

Lemma 16. *In the explore phase with beliefs supported on $[a_j, b_j]$, there always exists a price $p \geq a_j$ inducing the threshold t which satisfies $F_j(t) = \frac{1-\delta}{\delta}$.*

Proof. Note that the threshold equation for this stage can be rearranged as:

$$(t-a_j)F_j(t)\frac{\delta}{1-\delta} = (F_j(t)+1)(t-p).$$

Substituting in $F_j(t) = \frac{1-\delta}{\delta}$ and solving for p yields $p = t - \delta(t-a_j)$. □

To obtain our bound, we will assume the seller offers the following sequence of prices:

- If there exists some $p \in [a_j, b_j]$ inducing threshold p^*, offer p.
- Otherwise, offer a price which induces t satisfying $F_j(t) = \frac{1-\delta}{\delta}$.

We now argue that such a sequence of prices will eventually induce a threshold of p^*, if buyers' values are above p^*.

Lemma 17. *If both sellers have value at least p^*, then the above sequence of prices eventually induces threshold p^*.*

Proof. By Lemma 16, the seller will eventually reach a stage where the threshold t satisfying $F_j(t) = \frac{1-\delta}{\delta}$ is greater than p^*. We show that in this case, there is a price inducing a threshold $t = p^*$.

Assume the current beliefs for buyers who haven't rejected are distributed according to F_j with support $[a_j, b_j]$. Let t^* be the threshold for which $F_j(t^*) = \frac{1-\delta}{\delta}$, and assume $t^* > p^*$. Note that the threshold equation can be rearranged as:

$$\frac{t-p}{t-a_j} = \frac{F_j(t)}{F_j(t)+1}\frac{\delta}{1-\delta} \tag{5}$$

Consider substituting $t = p^*$ in the righthand side of (5). Since $t^* > p^*$, and since this function on the righthand side is increasing in t, we have:

$$0 < \frac{F_j(p^*)}{F_j(p^*)+1}\frac{\delta}{1-\delta} < \frac{F_j(t^*)}{F_j(t^*)+1}\frac{\delta}{1-\delta} < 1.$$

In other words, the righthand side of (5) for $t = p^*$ lies in the interval $[0, 1]$. Note that the lefthand side, with $t = p^*$, ranges from 0 at $p = p^*$ to 1 at $p = a_j$. Consequently, there is a price in $[a_j, p^*]$ which induces p^* as a threshold. □

We can now argue that under the above sequence of prices, the exploration phase will reach threshold p^* quickly if both agents have values above p^*. Formally:

Lemma 18. *Let $x = \frac{1-\delta}{\delta}$. If both sellers have value at least p^* and $F(p^*) \leq 1 - 1/e$ then after $1/x + 1$ rounds of the exploration phase using the price sequence defined above, we will have that the lower bound of the support is p^*.*

Proof. Let t_j be the threshold induced in the jth stage of the exploration phase, and assume $t_j < p^*$. From the above analysis, we have that $F(t_j) = 1 - (1-x)^j$. Set If we set $j = 1/x$, we obtain

$$F(t_j) = 1 - (1-x)^{1/x} \geq 1 - 1/e.$$

It therefore must be that after at most $1/x$ rounds, the exploration phase terminates with the threshold reaching p^*. In the subsequent round, the lower bound of the support will be the previous round's threshold, p^*. □

Lemma 19. *If $F(p^*) \leq 1-1/e$ then the equilibrium obtains revenue at least $\frac{1}{1-\delta}\frac{2}{3e}p^*(1-F(p^*))^2$.*

Proof. We lower bound the revenue with the revenue from the sequence of prices described above. The probability that both agents have values above p^*, and therefore that the threshold reaches p^*, is $(1 - F(p^*))^2$. By Lemma 18 the discount factor after reaching p^* is at most $\delta^{1+\frac{\delta}{1-\delta}} \geq \frac{2}{3e}$. After p^* is reached the seller prices the item at p^* for all remaining rounds and the item is accepted with probability one for a total of $\frac{1}{1-\delta}p^*$ revenue. Overall the revenue obtained by the seller with this price sequence is therefore at least $\frac{1}{1-\delta}\frac{2}{3e}p^*(1 - F(p^*))^2$. □

Proof of Theorem 15. Let OPT denote the total revenue from running the optimal auction for two buyers on F. Our benchmark for revenue is $\text{OPT}/(1-\delta)$. By concavity of the revenue we have that $OPT \leq 2p^*(1 - F(p^*))$. By Lemma 19 the equilibrium gets revenue

$$\frac{1}{1-\delta}\frac{2}{3e}p^*(1-F(p^*))^2 \quad \geq \quad \frac{\text{OPT}}{1-\delta}\frac{1}{3e}(1-F(p^*))$$

For distributions satisfying the monotone hazard rate assumption, it is a standard fact that $F(p^*) \geq 1 - 1/e$. We therefore have that the revenue of our equilibrium is at least $\frac{\text{OPT}}{1-\delta} \cdot \frac{1}{3e^2}$, proving the theorem. □

References

Battaglini, M. (2005). Long-term contracting with markovian consumers. *The American economic review*, 95(3):637–658.

Bulow, J. and Klemperer, P. (1994). Auctions vs. negotiations. Technical report, National Bureau of Economic Research.

Coase, R. H. (1972). Durability and monopoly. *The Journal of Law and Economics*, 15(1):143–149.

Conitzer, V., Taylor, C. R., and Wagman, L. (2012). Hide and seek: Costly consumer privacy in a market with repeat purchases. *Marketing Science*, 31(2):277–292.

Devanur, N. R., Peres, Y., and Sivan, B. (2015). Perfect bayesian equilibria in repeated sales. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 983–1002.

Fudenberg, D. and Tirole, J. (1983). Sequential bargaining with incomplete information. *The Review of Economic Studies*, 50(2):221–247.

Fudenberg, D. and Villas-Boas, J. M. (2006). Behavior-based price discrimination and customer recognition. *Handbook on economics and information systems*, 1:377–436.

Hart, O. D. and Tirole, J. (1988). Contract renegotiation and coasian dynamics. *The Review of Economic Studies*, 55(4):509–540.

Kennan, J. (2001). Repeated bargaining with persistent private information. *The Review of Economic Studies*, 68(4):719–755.

Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73.

Schmidt, K. M. (1993). Commitment through incomplete information in a simple repeated bargaining game. *Journal of Economic Theory*, 60:114–114.

A No-Learning Equilibrium

In this appendix, we give the full description of the no-learning equilibrium. The seller's strategy can be found in Algorithm 1, and the buyer's strategy in Algorithm 2. The beliefs which support this strategy profile are straightforward: if a buyer has ever accepted a positive price or rejected 0 (neither of which is on-path), the seller believes the buyer's value is deterministically the highest value in the initial support. Otherwise, the seller learns nothing about the buyer's value and offers the item for free every round. Moreover, it is clear from inspection that this equilibrium survives the simplicity refinement.

Algorithm 1: Zero-Revenue Equilibrium - Seller's Strategy

Input : Purchasing history h^k, initial belief support $[a, b]$.
Output: Price p_k

```
if Buyer has ever accepted a positive price then
    p_k = b;
else if Buyer has ever rejected a price of 0 then
    p_k = b;
else
    p_k = 0;
```

Algorithm 2: Zero-Revenue Equilibrium - Buyer's Strategy

Input : Purchasing history h^k, belief support $[a, b]$, value v, price p_k
Output: Purchasing decision for round k

```
if p_k = 0 then
    Accept;
else if p_k > 0 then
    if Buyer has ever accepted a positive price then
        Accept if and only if v ≥ p_k;
    else if Buyer has ever rejected a price of 0 then
        Accept if and only if v ≥ p_k;
    else
        Reject;
```

B Proof of Theorem 2

We will explicitly construct an equilibrium where the seller offers price p every round, no matter the buyer's action. We give the seller's strategy in Algorithm 3, and the buyers' strategy in Algorithm 4. Beliefs are straightforward - on-path, they are updated after the first buying decision and remain constant thereafter. If the seller has caused an off-path history by posting a price other than p, then they expect positive prices to be rejected for the rest of time, as in the zero-revenue equilibrium. As in the latter equilibrium, if a buyer accepts a positive price, then the seller assumes that they have the highest possible value for their distribution, and posts this value as the price until the end of time.

Algorithm 3: Folk Theorem Equilibrium - Seller's Strategy

Input : History h^k, Initial belief supports $[a_i, b_i]$ for all buyers i
Output: Price p_k

```
if p[k − 1] = (p, . . . , p) then
    p_k = p;
else if Any buyer has accepted a price other than p or 0 then
    p_k = b;
else
    p_k = 0
```

Algorithm 4: Folk Theorem Equilibrium - Buyer i's Strategy

Input : History h^k, Value v_i, Price p_k
Output: Purchasing decision for round k
if $\mathbf{p}[k] = (p, \ldots, p)$ **then**
  Accept if and only if $v_i \geq p_k$;
else if *Buyer i has accepted a positive price other than p or 0* **then**
  Accept if and only if $v_i \geq p_k$;
else if $p_k \neq 0$ **then**
  Reject;
else
  Accept;

To see that the strategies in Algorithms 3 and 4 are a PBE, we first argue that the seller is best responding. We consider the cases as they are stated in Algorithm 3:

- *All prices offered have been p:* In this case, buyers will behave as price takers if $p_k = p$, which might yield positive revenue. If the seller offers any other positive price, buyers will reject and demand the items for free for the rest of the game.
- *A price other than p has been offered and accepted by buyer i:* This is off-path. We may therefore set the seller's beliefs to be the highest value in the initial support, i.e. $v_i = b$. Moreover, according to the buyers' strategy, i will be a price-taker from now on. It follows that it is optimal for the seller to set a price of b.
- *A price other than p has been offered, but no buyer has accepted a positive price other than p:* In this case, buyers will only accept a price of 0, so the seller cannot get any utility with any price; they might as well post 0.

We now argue that an arbitrary buyer i is best responding, using the cases in Algorithm 4.

- *All prices offered, including p_k, have been p:* The seller will continue offering this price no matter what the buyer does. It follows that the buyer should accept if they could get positive utility from doing so.
- *Buyer i has accepted a positive price other than p or* 0*:* In this case, the seller believes that $v_i = b$, and will post price b forever. Buyer i should therefore reject, unless their value is b, in which case they weakly prefer to accept.
- *A price other than p has been offered, the only positve price accepted has been p, and $p_k \neq 0$.* Buyer i will get at most utility $v_i - p_k$ from accepting, as all future prices will be 1. Rejecting, meanwhile, will yield utility of $\frac{\delta}{(1-\delta)}$, as the seller will offer the items for free for all subsequent rounds, and all buyers will accept. If $\delta \geq 1/2$, then rejecting will be preferable for any choice of v_i and p_k.
- *A price other than p has been offered, the only positive price accepted has been p, and $p_k = 0$:* Rejecting will not change the seller's subsequent prices, and accepting will yield positive expected utility, so accepting is optimal.

C Alternate Sufficient Conditions for Eliminating Learning Equilibria

In this appendix, we give an alternate refinement which selects out all but equilibria in which the seller learns nothing about the buyer, and buyers accept every round. As in Section 4, we require that the seller posts prices which are above the bottom of the support of beliefs. Rather than upperbounding prices offered by the seller, we require that the buyers accept a price at the bottom of the support, whenever it is offered. Formally:

Definition 20. *A threshold PBE of the single-buyer game* respects lower bounds *if the following conditions hold for every history h^k with beliefs lower bounded by a:*

- *the buyer accepts any price at or below a.*
- *$\sigma_S^k(h^k) \geq a$.*

Theorem 21. *In the single-buyer game, let the value distribution F be supported on $[a, b]$, with $a > 0$. If $\delta > \frac{b}{a+b}$, then in any threshold PBE which respects lower bounds, the seller posts a every round, which is accepted by all buyers. In other words, no learning will occur.*

Note that Theorem 21 does not require that strategies be Markovian on path, though adding this requirement obviously does not change the result.

To prove the theorem, we show that any threshold PBE that respects lowerbounds must have a non-monotone cumulative allocation function around any threshold other than a or b. To simplify our analysis, we will consider only threshold PBE in which the threshold buyer for each round accepts in the first round. Moreover, since the seller never will offer a subsequent price below t, we may have the threshold type reject for the remainder of the game. For any threshold PBE, one can change the strategies of threshold buyers to an accept followed by nonstop rejection without violating equilibrium, as such a sequence of actions is one of their optimal choices. The change doesn't affect the seller's expected utility because such agents are a measure zero set.

Proof of Theorem 21. Consider a threshold PBE with value distribution supported on $[a, b]$. Assume the threshold buyer in the first round behaves as described above: they accept in the first round, and then reject in every subsequent round. In this equilibrium, we have $X(t) = 1$. We will show that there is a lower-valued agent with with total discounted allocation strictly greater than 1. This violates Lemma 7 unless $t = a$, proving the theorem.

We will first find a buyer with value less than t with high total discounted payments. To do this, note that after seeing a rejection in the first round, the seller could offer a for the rest of the game, which by the natural threhsolds assumption would yield revenue $a\frac{\delta}{1-\delta}$. Since the seller is best responding, this implies that $\mathbb{E}[P(v) \,|\, v \leq t] \geq a\frac{\delta}{1-\delta}$. Since $P(v)$ is increasing, it must be that there is a set of values with positive measure in $[a, t]$ with total discounted payments at least $a\frac{\delta}{1-\delta}$. Choose some v from this set. We have $P(v) \geq a\frac{\delta}{1-\delta}$.

We now lowerbound $X(v)$. To do this, note that the buyer could choose to reject every round, so $U(v) \geq 0$. This in turn implies that $vX(v) \geq P(v)$, and therefore that $X(v) \geq \frac{a}{v}\frac{\delta}{1-\delta}$. By our assumption that $\delta > \frac{b}{a+b}$, we have:

$$X(v) \geq \frac{a}{v}\frac{\delta}{1-\delta} \geq \frac{a}{b}\frac{\delta}{1-\delta} > 1.$$

This yields the desired non-monotonicity of $X(\cdot)$, contradicting Lemma 7. It must therefore be that $t = a$. The seller does not learn in such an equilibrium, as the same arguments will hold for every subsequent round. □

D Equilibrium for Digital Goods and Multiple Buyers

In this appendix, we describe a family of equilibria for $n \geq 2$ buyers with unlimited supply, and prove that the sellers and buyers are best responding. Our equilibria are parametrized by an integer $k \in \{1, \ldots, n\}$, which we call the *supply limit*, and a price p^t, which we call the *target price*. In rough terms, the seller starts with a low price, which gradually increases from round to round. Buyers drop out as the price grows too high. The seller stops raising the price when the number of buyers who remain is k or fewer, or when the price rises above p^t. Hence, the seller loosely implements a k-item ascending price auction over many rounds of sale.

We can more precisely describe our equilibrium by dividing it into an exploration phase and an exploitation phase. In round j of the exploration phase, the seller chooses a price p_j and threshold $t_j(p_j)$ which satisfy a threshold equation, described below. Buyers with value below the threshold reject p_j, and buyers with value above accept. The price and threshold rise from round to round, until eventually, one of three events occurs, each of which triggers the exploitation phase:

- Some number of buyers between 1 and the supply limit k accept p_k, while every other buyer rejects. The buyers who accepted all have value at least $t_j(p_j)$, so the seller posts this price for the rest of the game.

- All buyers reject p_k. Since this is the first round in which no buyer has accepted, there is some set of buyers who accepted the previous rounds price p_{j-1}. These buyers have value at least the previous round's threshold $t_{j-1}(p_{j-1})$. The seller posts this price for the rest of the game.

- More than k buyers accept p_j, but the corresponding threshold $t_j(p_j)$ exceeds the target price. The seller posts $t_j(p_j)$ for the rest of the game.

In the exploitation phase, the seller posts the same price p for the rest of the game. The exact price p is determined by the cases discussed above. The buyers whose values exceed p accept in every subsequent round. Those whose values do not exceed p reject in every subsequent round. Consequently, learning ceases; the seller instead extracts the lower bound of the support of the buyers with the highest such lower bound. This halt in price exploration is enforced with strategies that resemble those of the no-learning equilibrium of Theorem 5.

In the exploration phase, the prices offered each round and the buyers' responses are jointly governed by a threshold equation. Note that the exploration phase ends if all buyers reject the current price. Hence, in round j of the exploration phase, some set of buyers has accepted the previous round's price (and every previous price). Assume there are n_j such buyers. Let F_j denote the CDF of the common beliefs about these buyers' values, and let a_j denote the lower bound of this distribution's support. The current round's price p_j and threshold t_j must jointly satisfy the equation:

$$(t_j - p_j) = \frac{\delta}{1-\delta}(t_j - a_j) F_j(t_j)^{n_j - 1}. \tag{6}$$

The threshold equation states that a buyer with value t_j must be indifferent between accepting p_j and rejecting it. The lefthand side of (6) represents such a buyer's utility from accepting; they get utility $t_j - p_j$ from the current round, but do not earn any further utility, as all future prices will be at least t_j. The righthand side represents the threshold buyer's utility from rejecting. They receive no utility from the current round. Future prices will be at least t_j unless all other buyers also reject p_j, which occurs with probability $F_j(t_j)^{n_j-1}$. If this is the case, the seller offers the price a_j for the remaining rounds. The utility from accepting this price for the remainder of the game is $\frac{\delta}{1-\delta}(t_j - a_j)$.

Given a price p_j for round j, there may be be multiple choices of t_j satisfying the threshold equation (6). For any way of choosing such a threshold, there exists an equlibrium in which buyers' strategies all reflect this choice. To facilitate revenue analysis, we employ the following threshold selection rule:

- If the target price $p^t =$ satisfies $(p^t - p_j) = \frac{\delta}{1-\delta}(p^t - a_j)F_j(p^t)^{n_j-1}$, select $t_j = p^t$.
- Otherwise, choose the highest threshold t_j satisfying the threshold equation.

Every threshold selection rule determines a path for the posted price each round, which can be derived by recursively solving the seller's revenue maximization problem. In particular, let $t_a^b(p_j\,;\,n_j)$ be a threshold selection rule which depends only on the current price p_j, the number of buyers n_j who have not yet rejected, and the current beliefs, which are fully characterized in threshold equilibria by the limits of the support $[a_j, b_j]$ of buyers who have not yet rejected a price. The rule described above is one such selection rule. The seller chooses the current price p_j to maximize the following recursive expected revenue function:

$$
\begin{aligned}
R_{a_j}^{b_j}(p_j\,;\,n_j) \quad = \quad & F(t_{a_j}^{b_j}(p_j\,;\,n_j))^{n_j}\tfrac{\delta}{1-\delta}n_j a_j \\
& + \sum_{i=1}^{k}\binom{n_j}{i}(1 - F(t_{a_j}^{b_j}(p_j\,;\,n_j)))^i F(t_{a_j}^{b_j}(p_j\,;\,n_j))^{n_j-i}(ip_j + \tfrac{\delta}{1-\delta}it_{a_j}^{b_j}(p_j\,;\,n_j)) \\
& + \sum_{i=k+1}^{n_j}\binom{n_j}{i}(1 - F(t_{a_j}^{b_j}(p_j\,;\,n_j)))^i F(t_{a_j}^{b_j}(p_j\,;\,n_j))^{n_j-i}(ip_j + \delta R_{t_{a_j}^{b_j}(p_j\,;\,n_j)}^{b_j}(i)),
\end{aligned}
$$

where k is the supply limit parameter of the equilibrium, and $R_x^y(i)$ denotes the seller's optimal expected discounted revenue from a continuation in which there are i buyers who have not rejected, with beliefs supported on $[x, y]$.

The first term in the revenue recurrence is the revenue from the event that all buyers reject p_j. This occurs with probability $F(t_{a_j}^{b_j}(p_j\,;\,n_j))^{n_j}$ and causes the seller to post a_j for the remainder of the game. This price will be accepted by n_j buyers for $\frac{\delta}{1-\delta}$ discounted rounds. The first sum represents the events where some number $i \leq k$ buyers accepts p_j, which triggers the exploitation phase. The probability of this event for a given i is $\binom{n_j}{i}(1 - F(t_{a_j}^{b_j}(p_j\,;\,n_j)))^i F(t_{a_j}^{b_j}(p_j\,;\,n_j))^{n_j-i}$, and the revenue from this event is ip_j for round j's revenue, plus $\frac{\delta}{1-\delta}it_{a_j}^{b_j}(p_j\,;\,n_j)$ from selling to i buyers at price $t_{a_j}^{b_j}(p_j\,;\,n_j)$ for the rest of the game. Finally, the second sum accounts for the revenue from $i > k$ buyers accepting price p_j. In this case, the seller again makes ip_j revenue from round j, but additionally solves a smaller version of the same pricing problem on support $[t_{a_j}^{b_j}(p_j\,;\,n_j), b_j]$ with i buyers who have not yet rejected, yielding continuation revenue $\delta R_{t_{a_j}^{b_j}(p_j\,;\,n_j)}^{b_j}(i)$. Note that as long as $t_{a_j}^{b_j}(p_j\,;\,n_j)$ depends only on the current price p_j, number n_j

Algorithm 5: Seller's Strategy

Input : Purchasing history h^j, Support bounds $\{(a_i^j, b_i^j)\}_{i=1}^n$, First-round common support bounds (a, b).

Output: Price p_{j+1}

if $\exists\ \ell \leq j$ *and* i *such that* i *accepts* p_ℓ *but all types for* i *should reject* p_ℓ **then**

  Let ℓ^* be the most recent such ℓ;

  if $\exists m > \ell^*$ *such that* $p_m < b$ **then**

    $p_{j+1} = 0$;

  else

    $p_{j+1} = b$;

else

  $S_j = \text{argmax}_i\, a_i^j$;

  $a^j = \max_i a_i^j$;

  $b^j = \max_i b_i^j$;

  if k *or fewer agents accepted in round* j **or** $a \geq p^t$ **then**

    $p_{j+1} = a$;

  else

    $p_{k+1} = \arg\max_p R_{a^j}^{b^j}(p\,;\,|S_k|)$;

of buyers who have not yet rejected, and limits of the support $[a_j, b_j]$, then the price maximizing $R_{a_j}^{b_j}(p_j\,;\,n_j)$ will depend only on n_j, a_j, and b_j, and hence, will produce a Markovian equilibrium on path.

Having characterized the exploration and exploitation phases of our equilibrium, we now briefly describe the responses to off-path actions. As mentioned, the incentives in the exploitation phase resemble those of the zero-learning equilibrium of Theorem 5. We therefore describe only the exploration phase here. For any buyer i, rejecting of a price p_j and later accepting of a price $p_{j'} \geq p_j$ is incompatible with the strategy of any type for i. Consequently, the belief updates maybe arbitrary. After such an unexpected acceptance, the beliefs for buyer i are updated to a pointmass on the highest value in the support of the original distribution. The seller offers this high price for the rest of the game, preventing the deviating buyer from obtaining any future utility. The seller may deviate in the exploration phase by offering a price p_j for which there is no threshold t_j solving (6). In this case, all buyers reject p_j. The seller posts 0 for the remainder of the game, and the buyers reject any positive price, emulating the strategies of the no-learning equilibrium of Theorem 5.

A full description of the strategies and belief updates of the buyers and seller can be found in Algorithms 5, 6, and 8. Algorithm 7 describes the threshold selection rule we will use for revenue analysis.

Theorem 22. *The strategies and belief and belief updates described in Algorithms 5, 6, and 8 are a PBE as long as $\delta \geq 1/2$.*

We argue buyer and seller incentives separately, in Sections D.2 and D.3, respectively. Before performing this analysis, we will characterize the buyers' and seller's incentives from taking actions that deviate from the equilibrium path in Section D.1.

Algorithm 6: Buyer i's Strategy

Input : Purchasing history h^j, Support bounds $\{(a^j_{i'}, b^j_{i'})\}_{i'=1}^n$, value v_i, price p_{j+1}
Output: Purchasing decision for round $j+1$

if $\exists\ \ell \leq j$ *such that* i *accepted* p_ℓ *but all types for* i *should reject* p_ℓ **then**
    Let ℓ^* be the most recent such ℓ;
    if $p_m = b$ *for all* $\ell^* \leq m \leq j$ **then**
        Accept if and only if $p_{j+1} \geq v_i$;
    else
        Accept if and only if $p_{j+1} = 0$;
else if $\exists\ \ell \leq j$ *and* $i' \neq i$ *such that* i' *accepts* p_ℓ *but all types for* i' *should reject* p_ℓ **then**
    Accept if and only if $p_{j+1} = 0$;
else if $\exists \ell \leq j$ *such that* $p_\ell > p_{\ell+1}$ **and** $\mathbf{D}^\ell \neq R^n$ **then**
    Accept if and only if $p_{j+1} = 0$;
else
    for *all* $\ell \leq j$ **do**
        $S_\ell = \text{argmax}_{i'}\ a^\ell_{i'}$;
        $a^\ell = \max_{i'} a^\ell_{i'}$;
        $b^\ell = \max_{i'} b^\ell_{i'}$;
    if $\exists \ell \leq j$ *such that* $\mathbf{D}^\ell = R^n$ **or** $|S_\ell| \leq k$ **or** $a^\ell \geq p^t$ **then**
        if $\exists m > \ell$ *such that* $p_m \neq a^\ell$ **then**
            Accept if and only if $p_{j+1} = 0$;
        else
            Accept if and only if $p_{j+1} = a^j$;
    else if $i \notin S_j$ *or* $p_{j+1} < a^\ell$ **then**
        Accept if and only if $p_{j+1} = 0$;
    else
        Accept if and only if $v_i \geq t^{b_j}_{a^j}(p_{j+1}\,;\,|S_j|)$;

D.1 Off-Path Incentives

In this section, we show that for both the seller and the buyer, deviating from on-path behavior is costly; the seller sacrifices all revenue from the current round onward, and buyers sacrifice the opportunity to obtain utility in future rounds. Moreover, we will characterize the continuations derived from off-path histories, and show that their incentives are similar to the zero-learning equilibrium of Theorem 5.

Consider a history h^j up to round j in which at least one agent (either buyer or seller) has deviated from the equilibrium path. For the seller, the set of possible deviations consists of:

- Lowering the price in the exploration phase after a round in which one more more buyers accepts.
- Posting a price strictly below the support of the set of targeted buyers S_j in the exploration phase.
- Posting a price other than the bottom of the support a^j of the buyers being targeted in

Algorithm 7: Threshold Selection Rule

Input : Support bounds (a^j, b^j), price p_{j+1}, integer n^j
Output: Threshold $t_{a^j}^{b^j}(p_{j+1}\,;\,n^j)$
if $(p^t - p_{j+1}) = \frac{\delta}{1-\delta}(p^t - a^j)F_{a^j}^{b^j}(p^t)^{n_j - 1}$ **then**
  Return p^t;
else if $\exists t$ *such that* $(t - p_{j+1}) = \frac{\delta}{1-\delta}(t - a^j)F_{a^j}^{b^j}(t)^{n_j - 1}$ **then**
  Return $\max\{t\,:\,(t - p_{j+1}) = \frac{\delta}{1-\delta}(t - a^j)F_{a^j}^{b^j}(t)^{n_j - 1}\}$;
else
  Return ∞;

the exploitatation phase.

- Posting a price other than the top of the initial support b after a buyer has taken an off-path action.

Note that by the definition of the buyers' strategies in Algorithm 6, buyers will respond to a deviation by the seller by entering a continuation in which they adopt the strategies from the zero-learning (and zero-revenue) equilibrium of Theorem 5 (which are best responses as long as $\delta \geq 1/2$). Hence, the seller cannot obtain any revenue after taking an off-path action, and any prices in a continuation in which the seller is the agent to most recently deviate will match the zero-learning equilibrium.

Now consider the consequences of an off-path action by one or more buyers. Buyer actions are off-path when the buyer accepts a price that all types for that buyer would reject (according to the equilibrium strategies and the current beliefs). As long as the seller has not taken one of the aforementioned deviations after the off-path buyer actions, the seller strategy dictates that they offer a high price of b, which is a best response because the deviating buyers become price-takers (and beliefs update for those buyers to pointmasses on b). Consequently, any buyer who deviates from the equilibrium path will not attain any utility in future rounds. We summarize these conclusions below:

Lemma 23. *Consider a history h^j in which one or more agents takes an off-path action. If the seller is the most recent agent to take such an action, the seller posts a price of* 0 *for every subsequent round. If one or more buyers are the most recent agents to take an off-path action by accepting a price which all types would reject according to the current beliefs, the seller posts a price of b for every subsequent round (which is rejected by the buyers). These strategies are best responses as long as $\delta \geq 1/2$.*

Corollary 24. *If the seller takes an off-path action, they earn no revenue in future rounds.*

Corollary 25. *If a buyer takes an off-path action, they earn no utility in future rounds.*

Corollaries 24 and 25 will facilitate the on-path analysis in Sections D.2 and D.3.

D.2 Buyer Incentives

We now show that in any on-path history, it is a best response to play the strategy described in Algorithm 6. This implies that the buyers are best responding. We analyze the exploration phase and exploitation phases separately.

Algorithm 8: Belief updates

Input : Purchasing history h^{j+1}, Current support bounds $\{(a_i^j, b_i^j)\}_{i=1}^n$, First-round common support bounds (a, b).

Output: Updated support bounds $\{(a_i^{j+1}, b_i^{j+1})\}_{i=1}^n$.

```
S_j = argmax_i a_i^j;
for i ∈ {1, ..., n} do
    if all types for i should reject p_{j+1} then
        if buyer i rejected p_{j+1} then
            a_i^{j+1} = a_i^j, b_i^{j+1} = b_i^j
        else
            a_i^{j+1} = b, b_i^{j+1} = b
    else if all types should accept then
        if buyer i accepted then
            a_i^{j+1} = a_i^j, b_i^{j+1} = b_i^j
        else
            a_i^{j+1} = b, b_i^{j+1} = b
    else
        if buyer i accepted then
            a_i^{j+1} = t_{a^j}^{b^j}(p_{j+1}, |S_j|), b_i^{j+1} = b_i^j
        else
            a_i^{j+1} = a_i^j, b_i^{j+1} = t_{a^j}^{b^j}(p_{j+1}, |S_j|)
```

Exploration Phase In the exploration phase, the seller posts a price which increases from round to round. Buyers who have not yet rejected play according to a nontrivial threshold strategy given by the threshold equation (6). Buyers who have rejected in the past continue to reject the current price, as it is above their value. For these latter buyers, accepting a price after rejecting would cause the seller to offer the top of the initial support, b, for the remaining rounds, so there is no utility to be gained from any deviation. For the buyers who play according to equation (6), we need to show that buyers with value above the threshold t_j prefer to accept, and those with value below t_j prefer to reject. We argue from the perspective of an arbitrary buyer i who has not yet rejected a price in the exploration phase (and is therefore being "targeted" by the seller), facing a price p_j for round j. We assume the current support of beliefs for i and all other buyers who are being targeted is $[a_j, b_j]$, and that there are n_j such buyers. We denote the CDF for the current beliefs of these buyers by $F_j(\cdot)$.

Assume $v_i < t_j$. If buyer i accepts p_j, all subsequent prices will be at least t_j. Hence, the utility from accepting, u_A, is upper bounded as $u_A \leq (v_i - p_j)$. A lower bound on the utility from rejecting, u_R, comes from noting that with probability $F(t_j)^{n_j-1}$, all other targeted buyers will reject as well. In this event, the seller posts a price of a_j for the rest of the game, yielding a utility of $F(t_j)^{n_j-1}(v_j - a_j)\frac{\delta}{1-\delta}$. The utility difference between rejecting and accepting therefore satisfies:

$$u_R - u_A \geq F(t_j)^{n_j-1}(v_i - a_j)\tfrac{\delta}{1-\delta} - (v_i - p_j).$$

Note that because the seller is offering an on-path price, $p_j \geq a_j$, so $t_j - p_j \leq t_j - a_j$, and hence the threshold equation (6) implies that $F_j(t_j)^{n^j-1}\frac{\delta}{1-\delta} \leq 1$. Writing $x = t_j - v_i$, we have:

$$u_R - u_A \geq F(t_j)^{n_j-1}(t_j - x - a_j)\tfrac{\delta}{1-\delta} - (t_j - x - p_j). \tag{7}$$

Applying the threshold equation (6), we can rearrange (7) as

$$u_R - u_A \geq x(1 - F(t_j)^{n_j-1}\tfrac{\delta}{1-\delta}) \geq 0,$$

which implies that rejecting is optimal.

Now assume $v_i \geq t_j$. We must prove that it is optimal for buyer i to accept, i.e. $u_A \geq u_R$. We first lower bound u_A by noting that buyer i gets utility $v_i - p_j$ from accepting in round j, and from the event that all other buyers reject an additional $F_j(t_j)^{n_j-1}(v_i - t_j)\frac{\delta}{1-\delta}$, as the seller will post price t_j for the rest of the game, which buyer i may accept. In total, we have that u_A is at least $(v_i - p_j) + F_j(t_j)^{n_j-1}(v_i - t_j)\frac{\delta}{1-\delta}$.

To upper bound u_R, we consider two events: the event that every other buyer also rejects, and the event that at least one other buyer accepts. In the former event, which occurs with probability $F_j(t_j)^{n_j-1}$, the seller posts a_j for the rest of time. The best buyer i can do is to accept this price, yielding a utility contribution of $F_j(t_j)^{n_j-1}(v_i - a_j)\frac{\delta}{1-\delta}$. In the latter event, which occurs with probability $1 - F_j(t_j)^{n_j-1}$, the seller may accept in one later round. This acceptance causes i's beliefs to update to a pointmass on b, and leads the seller to post this price every subsequent round. Consequently, buyer i may only accept for one round after rejecting. The maximum utility attainable from such a purchase is $\delta(v_i - t_j)$, as the discount factor will always be at most δ and the price at least t_j. Hence, we have an upper bound of $F_j(t_j)^{n_j-1}(v_i - a_j)\frac{\delta}{1-\delta} + \delta(1 - F_j(t_j)^{n_j-1})(v_i - t_j)$. Combining the upper and lower bounds, we obtain:

$$u_A - u_R \geq (v_i - p_j) + F_j(t_j)^{n_j-1}(v_i - t_j)\tfrac{\delta}{1-\delta} - F_j(t_j)^{n_j-1}(v_i - a_j)\tfrac{\delta}{1-\delta} - \delta(1 - F_j(t_j)^{n_j-1})(v_i - t_j)$$

Writing $x = v_i - t_j$ and applying the threshold equation, we obtain:

$$\begin{aligned} u_A - u_R &\geq x(1 + F_j(t_j)^{n_j-1}\tfrac{\delta}{1-\delta} - F_j(t_j)^{n_j-1}\tfrac{\delta}{1-\delta} - \delta(1 - F_j(t_j)^{n_j-1})) \\ &= x(1 - \delta(1 - F_j(t_j)^{n_j-1}) \geq 0. \end{aligned}$$

Consequently, buyer i prefers to accept p_j.

Exploitation Phase In the exploitation phase, there are two classes of buyers: those being targeted and those who dropped out at an earlier price. If any number of buyers accepted the price in the round that triggered the exploitation phase, those buyers are the targeted set. If all buyers rejected, then the targeted set is the set of buyers who accepted in the previous round. All other buyers dropped out in a prior round.

For the buyers who dropped out, the current price is above their value, and will remain that way for the remainder of the game. They consequently are indifferent between all actions, including those specified by the equilibrium strategy. The buyers being targeted will be offered a price at the bottom of the support of the current beliefs on their values. This price is unresponsive to their action, so they would prefer to accept every round.

D.3 Seller Incentives

We break the seller's incentives into the exploration phase and the exploitation phase as usual. In the exploration phase, the choice is between offering an on-path price with a threshold response and an off-path price without. An off-path price will be rejected, and buyers will initiate continuation strategy in which they reject all positive prices yielding no revenue. It is therefore optimal for the seller to offer a price with a threshold response each round. In the exploitation phase, the in buyers being targeted will only accept the lower bound of the support of the beliefs on their values (and any other price will initiate a zero-revenue response from the buyers), so offering the lower bound is clearly optimal.

E Proof of Theorem 10

In this appendix, we prove Theorem 10: there exists a supply limit and target price for the digital goods equilibrium such that the seller obtains revenue which is a constant approximation to the optimal single-shot revenue for the digital goods setting per round. The proof mirrors that of the revenue result for limited supply in Appendix **??**. We give a pricing policy for the seller which is supported by threshold responses, and achieves the desired revenue guarantee. Since the seller chooses the revenue-optimal such pricing policy, this revenue is a lower bound on the seller's revenue.

Our choice of supply limit will depend on the quantile q^* of the monopoly price $p^* = \max_p(1-F(p))p$. Recall that we define the quantile of p^* as $q^* = 1 - F(p^*)$. By assumption, $q^* \geq 1/n$. For $q^* \in [1/n, 2/(n+1)]$, we will use a supply limit of $k^* = 1$. For $q^* \in [2/(n+1), 1]$, we use $k^* = \min(\lfloor q^*(n+1) \rfloor - 1, n-1)$. The constant-approximation result is robust to the choice of supply limit, in the sense that any supply limit that yields a comparable ex ante probability of sale to the optimal mechanism will suffice. The particular choice of k^* we adopt is for ease of technical exposition.

The pricing policy we use as a revenue benchmark will post prices with high thresholds in an attempt to learn aggressively and reach the target price as quickly as possible. The pricing policy will target one of two prices, again selected for ease of analysis. If $q^* \geq 2/(n+1)$, our pricing policy will target the price with quantile $\min(\lfloor q^*(n+1) \rfloor/(n+1), n/(n+1))$. If $q^* \in [1/n, 2/(n+1))$, we will instead target quantile $2/(n+1)$. For regular distributions, offering prices close to the monopoly price produces similar revenue, so this will suffice to prove a constant approximation, assuming the price target is reached quickly. We will also prove this latter fact, which will imply the theorem.

We first describe the pricing policy. Let p^t denote the target price described previously. After each round j, there is a set S_j of buyers who have either accepted in the previous round, or if all buyers rejected in the previous round, then accepted in the round before that. Let $n_j = |S_j|$, and let F_j denote the common CDF of the beliefs of the agents in S_j, with support $[a_j, b_j]$. The pricing policy is:

In Section E.1, we prove that the policy in Algorithm 9 is valid: in each round where more than k^* buyers have yet to reject, either the seller can post a price with threshold equal to p^t or there exists a price which induces a threshold t with $F_j(t) = \sqrt[n_j-1]{\frac{1-\delta}{\delta}}$. Consequently, the policy will reach the price p^t as long as at least k^*+1 buyers have value above p^t. In Section E.2, we then prove that this pricing policy attains a constant fraction of the revenue that would result from posting p^* to each agent each round (i.e. the Myerson revenue-mechanism for the initial

Algorithm 9: Pricing policy to lower bound equilibrium revenue.

if $F_j(p^t) = 0$ **then**
  Post p^t.
else if k^* *or fewer buyers accepted last round* **then**
  Post a_j.
else if *There is a price* p *with threshold* $t(p) = p^t$ **then**
  Post p.
else
  Post a price p which induces threshold $t(p)$ satisfying $F_j(t(p)) = \sqrt[n_j-1]{\frac{1-\delta}{\delta}}$.

distribution).

E.1 Validity of Pricing Policy

We now prove that the pricing policy of Algorithm 9 is valid. Either one of the three conditions in the algorithm holds, or there exists a price p which induces a corresponding threshold $t(p)$ such that $F_j(t(p)) = \sqrt[n_j-1]{\frac{1-\delta}{\delta}}$. This allows us to characterize the price path under this policy. As long as more than k^* agents have accepted the previous price and it is not possible to induce p^t as a threshold, the seller learns aggressively. If k^* or fewer buyers accept, the seller stops learning and exploits the highest-valued buyers at the bottom a_j of the support of their beliefs. Otherwise, the pricing policy will reach the target price of p^t, and the seller will post this price for the remainder of the game.

Lemma 26. *Assume at least k^*+1 buyers have not yet rejected by round j. Let t^* be the unique value satisfying $F_j(t^*) = \sqrt[n_j-1]{\frac{1-\delta}{\delta}}$. Then there is a price with threshold response t^*.*

For the limited supply setting, the result followed from letting the price p range over the possible values it might take and noting the continuity of the accept side of the threshold equation in p. For digital goods, the argument is even simpler:

Proof. Set $p = a_j$. We have that the threshold equation becomes $t^* - a_j = (t^* - a_j)F(t^*)^{n_j-1}\frac{\delta}{1-\delta}$. Substituting $F_j(t^*) = \sqrt[n_j-1]{\frac{1-\delta}{\delta}}$ yields the result. □

Lemma 26 ensures that the seller is able to learn aggressively as long as the number of remaining buyers exceeds the supply limit. The next lemma shows that once the seller has learned enough, they can halt exploration and induce p^t as a threshold. This ensures that the pricing policy of Algorithm 9 eventually terminates with the seller posting p^t, as long as k^* or more buyers have value which exceeds this target price.

Lemma 27. *Assume at least $k^* + 1$ buyers have not yet rejected by round j. Let t^* be the unique value satisfying $F_j(t^*) = \sqrt[n_j-1]{\frac{1-\delta}{\delta}}$. If $t^* \geq p^t$, then there is a price which induces p^t as a threshold.*

Proof. We must show that as long as $t^* > p^t$, then there is a price p such that

$$(p^t - p) = (p^t - a_j)F_j(p^t)^{n_j-1}\frac{\delta}{1-\delta}.$$

We may rearrange this as

$$\frac{p^t - p}{p^t - a} = F_j(p^t)^{n_j - 1} \tfrac{\delta}{1-\delta}. \tag{8}$$

We first show that the righthand side lies in $[0, 1]$. This follows from noting that

$$1 = F_j(t^*)^{n_j-1} \tfrac{\delta}{1-\delta} \geq F_j(p^t)^{n_j-1} \tfrac{\delta}{1-\delta} \geq 0,$$

since $F_j(t^*) = \sqrt[n_j-1]{\frac{1-\delta}{\delta}}$ and $t^* \geq p^t$. We then note that we may choose p to make the lefthand side of (8) range over all of $[0, 1]$ by choosing $p \in [a_j, p^t]$. It follows that there exists a choice of p that satisfies (8). □

E.2 Revenue Guarantee for Pricing Policy

We now argue that the expected revenue attained by the seller were they to use the pricing policy in Algorithm 9 is a constant fraction of the Myerson benchark, which is $n(1 - F(p^*))p^*/(1 - \delta)$, i.e. the revenue from selling to the agents each round using the revernue-optimal single-shot mechanism for the initial value distribution. The revenue of our pricing policy is lower bounded by

$$(k^* + 1)p^t \Pr[\mathcal{E}_{\geq k^*+1}] \frac{\mathbb{E}[\delta^{j^*} \mid \mathcal{E}_{\geq k^*+1}]}{1 - \delta},$$

where j^* is the round in which the seller first offers p^t and $\mathcal{E}_{\geq k^*+1}$ is the event that at least $k^* + 1$ buyers have value at least p^t (or more generally, we define $\mathcal{E}_{\geq k}$ to be the event that at least k buyers have value at least p^t).

The argument consists of two main steps. The first step is to show that $(k^* + 1)p^t \Pr[\mathcal{E}_{\geq k^*+1}]$ is within a constant factor of the single-shot Myerson revenue $np^*(1 - F(p^*))$. This will follow from basic probability theory and regularity of the value distributions. The second step is to prove that $\mathbb{E}[\delta^{j^*} \mid \mathcal{E}_{\geq k^*+1}]$ is bounded below by a constant. In other words, the pricing policy in Algorithm 9 converges to the price p^t quickly, conditioned on sufficiently many buyers having a high value. This will require a somewhat involved analysis of the stochastic process that governs the prices's path through buyers' quantile space.

Single-Shot Revenue Approximation We now show that $(k^* + 1)p^t \Pr[\mathcal{E}_{\geq k^*+1}]$, which can be interpreted as the probability of our pricing policy reaching p^t conditioned on $\mathcal{E}_{\geq k^*+1}$, multiplied by the revenue from a single round of sale to $k^* + 1$ buyers, is a constant approximation to the Myerson single-shot revenue $np^*(1 - F(p^*))$. We argue two cases, based on whether $q^* \in [1/n, 2/(n + 1)]$ or $q^* \geq 2/(n + 1)$. We begin with the former case.

Lemma 28. *Assume $q^* \in [1/n, 2/(n + 1)]$. Then $Pr[\mathcal{E}_{\geq k^*+1}] \geq 8n(1 - F(p^*))$.*

Proof. The probability that at least two buyers have quantile at most $2/(n + 1)$ is at least the probability that at least two buyers have quantile at most $1/n$. We may write this probability as

$$1 - \left[\left(1 - \frac{1}{n}\right)^n + \left(1 - \frac{1}{n}\right)^{n-1} \right].$$

This function is increasing in n, and hence it is lowerbounded by its value at $n = 2$, with value $1/4$. Since $1 - F(p^*) \leq 2/(n + 1)$, we also have that $n(1 - F(p^*)) \leq 2n/(n + 1) \leq 2$. Combining these two bounds yields the lemma. □

Lemma 29. *Assume $q^* \in [1/n, 2/(n+1)]$. Then $p^t \geq p^*/2$.*

Proof. By assumption, the buyers' initial value distribution is regular. By definition, this means the price-posting revenue curve $R(q) = qF^{-1}(1-q)$ is concave. We may therefore write:

$$\frac{2p^t}{n+1} = R\left(\frac{2}{n+1}\right) \geq \frac{1-\frac{2}{n+1}}{1-\frac{1}{n}} R(1/n) = \frac{1-\frac{2}{n+1}}{1-\frac{1}{n}} \cdot \frac{p^*}{n}.$$

Rearranging, we obtain

$$p^t \geq \frac{n+1}{2n} \cdot \frac{1-\frac{2}{n+1}}{1-\frac{1}{n}} p^* = p^*/2.$$

□

Combining Lemmas 28 and 29 yields a 16-approximation when $q^* \in [1/n, 2/(n+1)]$. We now turn to the case where $q^* \geq 2/(n+1)$.

Lemma 30. *Assume $q^* \geq 2/(n+1)$. Then $Pr[\mathcal{E}_{\geq k^*+1}]$ is lower bounded by a constant.*

Proof. Let q^t denote $1 - F(p^t)$. Note that $\Pr[\mathcal{E}_{\geq k^*+1}]$ is equal to the probability of $\min(\lfloor (n+1)q^* \rfloor, n-1) = (n+1)q^t$ successes among n Bernoulli trials with weight q^t. More abstractly, denote by $P(k, n, p)$ the probability of at least k successes from n Bernoulli trials of weight p. Fixing k, note that $P(k, n, p)$ is increasing in p. For our problem, $P(\min(\lfloor (n+1)q^* \rfloor, n-1), n, q^t)$ will therefore be minimized at integral values of $\lfloor (n+1)q^* \rfloor$. We may thus reduce the problem to lowerbounding $P(k, n, k/(n+1))$ for k integral.

To derive a crude lower bound on $P(k, n, k/(n+1))$, we use the Chernoff-Hoeffding inequality, which states that the probability of $k-1$ or fewer successes is at most $e^{-\frac{1}{2}(1-(k-1)k\frac{n}{n+1})^2 \frac{kn}{(n+1)}}$. Upperbounding this quantity amounts to analyzing the integral minima of the exponent, which can be shown to be at $k = 2$ for all values of n, and which are bounded away from 0, implying that the probability of $k-1$ or fewer successes is bounded away from 1. □

Lemma 31. *Assume $q^* \geq 2/(n+1)$. Then $k^* + 1 \geq \frac{2}{3}n(1 - F(p^*))$.*

Proof. If $q^* = 1$, then $n(1 - F(p^*)) = n$ and $k^* + 1 = n$ as well. Hence, we consider $q^* < 1$. In this case, we have that $q^* \in [j/(n+1), (j+1)/(n+1))$ for some $j \in \{2, \ldots, n\}$. We may upper bound $n(1 - F(p^*))$ by $\frac{n}{n+1}(j+1) \leq j+1$. We may lower bound $k^* + 1$ by j. Since $j \geq 2$, this implies the desired inequality. □

Combining Lemmas 30 and 31 with the fact that since $q^t \leq q^*$, $p^t \geq p^*$ yields a constant-approximation for the case where $q^* \geq 2/(n+1)$, and thus overall.

Speed of Convergence We now prove that $\mathbb{E}[\delta^{j^*} \mid \mathcal{E}_{\geq k^*+1}]$ is lowerbounded by a constant. The stochastic process we must analyze is the following: each round j, there are n_j buyers who have not yet rejected a price, and the current beliefs for these buyers has CDF F_j and support $[a_j, b_j]$. The seller posts a price p_j which induces a threshold t_j satisfying $F_j(t_j)^{n_j-1} = \frac{1-\delta}{\delta}$. In other words, the price p_j eliminates the bottom $(1 - \sqrt[n_j-1]{\frac{1-\delta}{\delta}})$-fraction of the support of the current beliefs, with respect to the measure F_j. In what follows, we call $(1 - \sqrt[n_j-1]{\frac{1-\delta}{\delta}})$ the *learning rate* for n_j buyers, which we denote r_{n_j}.

Let let $\mathcal{E}_k$ denote the probability that *exactly* k buyers have value above p^t. For any such event, we may construct a continuous upper bound on j^* conditioned on $\mathcal{E}_k$ in the following way:

Lemma 32. *For any $k \geq k^* + 1$, conditioned on $\mathcal{E}_k$, the following inequality holds:*

$$j^* \leq \log_{1-r_k} \frac{q^t}{q_{k+1}} + \sum_{i=k+1}^{n-1} \log_{1-r_i} \frac{q_i}{q_{i+1}} + \log_{1-r_n} q_n + n \tag{9}$$

where $q^t = 1 - F(p^t)$ and q_i denotes the quantile of the ith highest-valued agent.

Proof. In the event $\mathcal{E}_k$, exactly k buyers have quantile less than $2/(n+1)$. Until the first buyer rejects a price, the seller posts a price with threshold t_j satisfying $F(t_j) = r_n$. After j rounds of such prices, the bottom of the support of F_j will have quantile $(1 - r_n)^j$. Consequently, it will take $\lceil \log_{1-r_n} q_n \rceil \leq \log_{1-r_n} q_n + 1$ rounds to eliminate the first buyer.

We can make a similar argument to bound the time it takes to eliminate the buyer with the $i-1$st highest value (i.e. the buyer with quantile q_{i-1}) given that the ith highest-valued buyer was eliminated in the previous round. If the buyer with the ith highest value (henceforth, "buyer i"), was eliminated at the same time as buyer $i-1$ (i.e. both buyers rejected for the first time in the previous round), then for the sake of the argument, we may bound the number of rounds to eliminate buyer $i-1$ trivially by 1. Otherwise, the quantile of the bottom of the support of the beliefs of the buyers who have not rejected lies in $[q_{i-1}, q_i]$. Hence, the number of additional rounds required to eliminate buyer $i-1$ is at least $\lceil \log_{1-r_i} q_i/q_{i+1} \rceil \log_{1-r_i} q_i/q_{i+1} + 1$.

Finally, we may make a similar argument to argue that the time required for the price reach p^* is at most $\lceil \log_{1-r_k} \frac{2/(n+1)}{q_{k+1}} \rceil \leq \log_{1-r_k} \frac{2/(n+1)}{q_{k+1}} + 1$. Noting that j^* is the sum of the time required to eliminate each subsequent buyer and then raise the price to p^* once all but k have been eliminated, and summing the bounds for each of these times yields the righthand side of (9). □

Further applying Jensen's inequality to the upper bound in (9), we may replace each quantile with its expectation. Noting that this new bound is worst when $k = k^* + 1$, we obtain

Lemma 33. $\mathbb{E}[\delta^{j^*} \mid \mathcal{E}_{\geq k^*+1}] \geq \delta^{\sum_{i=2}^{n} \log_{1-r_i} \frac{i}{i+1} + n}$.

Proof. Note that δ^{j^*} is a convex function of j^*. Hence, we may apply Jensen's inequality to obtain $\mathbb{E}[\delta^{j^*} \mid \mathcal{E}_k] \geq \delta^{\mathbb{E}[j^* \mid \mathcal{E}_k]}$.

The bound of Lemma 32 implies

$$\mathbb{E}[j^* \mid \mathcal{E}_k] \leq \mathbb{E}\left[\log_{1-r_k} \frac{2/(n+1)}{q_{k+1}} + \sum_{i=k+1}^{n-1} \log_{1-r_i} \frac{q_i}{q_{i+1}} + \log_{1-r_n} q_n + n \mid \mathcal{E}_k\right], \tag{10}$$

where the expectation is over the quantiles $q_{k+1}, \ldots, q_n$, which are distributed as the order statistics of $n-k$ uniform draws from the interval $[2/(n+1), 1]$. We may rearrange the righthand side of (10) as:

$$\mathbb{E}\left[\sum_{i=k+1}^{n} \log q_i \left(\frac{1}{\log(1-r_i)} - \frac{1}{\log(1-r_{i-1})}\right) + \frac{\log \frac{2}{n+1}}{\log(1-r_k)} + n \mid \mathcal{E}_k\right]. \tag{11}$$

Note that $\frac{1}{\log(1-r_i)} - \frac{1}{\log(1-r_{i-1})}$ is positive for all i. Consequently, Jensen's inequality implies that we may upper bound the expectation of the function in the lefthand side of (10) by the function of the expectation of the q_is. We therefore have:

$$\mathbb{E}[j^* \mid \mathcal{E}_k] \leq \log_{1-r_k} \frac{2/(n+1)}{\mathbb{E}[q_{k+1} \mid \mathcal{E}_k]} + \sum_{i=k+1}^{n-1} \log_{1-r_i} \frac{\mathbb{E}[q_i \mid \mathcal{E}_k]}{\mathbb{E}[q_{i+1} \mid \mathcal{E}_k]} + \log_{1-r_n} \mathbb{E}[q_n \mid \mathcal{E}_k] + n. \tag{12}$$

Since $q_{k+1}, \ldots, q_n$ are distributed as the order statistics of $n-k$ uniform draws from the interval $[2/(n+1), 1]$, we may have $\mathbb{E}[q_i \mid \mathcal{E}_k] = \frac{i-k}{n-k+1}\left(1 - \frac{2}{n+1}\right) + \frac{2}{n+1}$. Substituting these quantities yields

$$\mathbb{E}[j^* \mid \mathcal{E}_k] \leq \log_{1-r_k} \frac{2/(n+1)}{\frac{k+1-k}{n-k+1}\left(1 - \frac{2}{n+1}\right) + \frac{2}{n+1}} + \sum_{i=k+1}^{n-1} \log_{1-r_i} \frac{\frac{i-k}{n-k+1}\left(1 - \frac{2}{n+1}\right) + \frac{2}{n+1}}{\frac{i+1-k}{n-k+1}\left(1 - \frac{2}{n+1}\right) + \frac{2}{n+1}}$$
$$+ \log_{1-r_n}\left(\frac{n-k}{n-k+1}\left(1 - \frac{2}{n+1}\right) + \frac{2}{n+1}\right) + n. \tag{13}$$

One can verify that this quantity is largest for $k = 2$, regardless of δ and n. Conditioned on the event $\mathcal{E}_2$, we have that $\mathbb{E}[q_i] = i/n + 1$. For all $k \geq 2$, we may therefore simplify the righthand side of (12) as

$$\mathbb{E}[j^* \mid \mathcal{E}_k] \leq \sum_{i=2}^{n} \log_{1-r_i} \frac{i}{i+1} + n. \tag{14}$$

□

Lemma 34. *The quantity $\delta^{\sum_{i=2}^n \log_{1-r_i} \frac{i}{i+1} + n}$ is lowerbounded by a constant for $\delta \geq n/(n+1)$.*

Proof. We may write

$$\delta^{\sum_{i=2}^n \log_{1-r_i} \frac{i}{i+1} + n} = \delta^n \prod_{i=2}^{n} \delta^{\log_{1-r_i} \frac{i}{i+1}}.$$

Note that for all i, the function $\delta^{\log_{1-\sqrt[i-1]{(1-\delta)/\delta}} \frac{i}{i+1}}$ is increasing in δ, so we may replace δ by its lower bound.

We next show that $\log_{1-\sqrt[i-1]{1/n}} \frac{i}{i+1}$ is at most $n^{1/(i-1)}$. To see this, note that $(1 - \sqrt[i-1]{1/n})^x$ is decreasing in x, taking value 1 at $x = 0$ and taking a value between $1/4$ and $1/e$ at $n^{1/(i-1)}$. For $i \geq 2$, $i/(i+1) \geq 1/e$, and hence the value of x that solves $(1 - \sqrt[i-1]{1/n})^x = i/(i+1)$ must be at most $n^{1/(i-1)}$. The lemma then follows from noting that $\sum_{i=1}^n n^{1/i}$ is $O(n)$. □

F Proof of Theorem 14

In this appendix, we prove a generalization of Theorem 5, stating that in any threshold equilibrium that satisfies the no-backtracking and natural prices refinements, the seller does not learn about buyers' values, and instead posts at the bottom of buyers' supports. Our proof generalizes that of Theorem 5; we show that the existence of a non-trivial threshold for some buyer implies a non-monotonicity in that buyer's expected total discounted allocation. This contradicts the monotonicity requirement of Theorem 8.

We begin by stating a version of Theorem 8 which holds in a setting with many buyers. The key difference is that in this new setting, randomness over other buyers' types might create uncertainty over the price in future rounds, so monotonicity of allocations holds with respect to *expected* allocations. We begin by defining expected discounted allocations.

Definition 35. *Given a PBE of the digital goods discriminatory pricing game with fixed value distributions, let $x_i^j(\mathbf{v})$ be an indicator variable of whether or not buyer i purchases in round j under the actions generated by the value profile $\mathbf{v}$. The* expected discounted allocation *for agent i with value v_i is given by:*

$$X_i(v_i) = \sum_{j=1}^{\infty} \delta^{j-1} \mathbb{E}_{\mathbf{v}_{-i}}[x_i^j(\mathbf{v})]$$

Any buyer i with value v_i may simulate the actions of some other value v_i' and obtain the expected discounted allocations and payments of that other value. Consequently, the proof of monotonicity condition of Theorem 8 applies. We obtain:

Lemma 36. *In any PBE for the digital goods setting with discriminatory pricing, for any buyer i, the expected total discounted allocation $X_i(v_i)$ is nondecreasing in v_i.*

Having established this fact, we now show that if there is a round j in which the seller post a price p_j and there is some buyer i who accepts with positive probability and rejects with positive probability, i.e. the seller learns, then as long as $\delta > 1/2$, there must be a nonmonotonicity in i's expected total discounted allocations.

Consider a history in which there is some round j in which some buyer i has a non-trivial threshold response to the price p_i^j. That is, their threshold t_i^j lies in the support $[a_i^j, b_i^j)$ of the current beliefs for their value. Since PBE requires that continuation strategies from every history be best responses for all players, we may assume for simplicity that j is the first round, and we suppress the round index j in what follows. As we did to prove Theorem 5, we note that all that is required to produce a contradiction is a type $v_i' < t_i$ such that $X_i(v_i') > 1$. If this occurs, we may break ties such that when buyer i has value t_i, they accept in the first round, and never accept a subsequent price without changing the incentives in equilibrium, as the threshold type has measure zero. In this new equilibrium, we have $X_i(t_i) = 1$. This yields the following lemma:

Lemma 37. *In any PBE for the digital goods setting with discriminatory pricing and buyer i, let t_i be the threshold type for buyer i under the first round's price. Then for any value $v_i \leq t_i$, $X_i(v_i) \leq 1$.*

We next show that if buyer i has a non-trivial threshold response in the first round, then there must be a value $v_i \leq t_i$ such that $X_i(v_i) > 1$, which will prove the theorem. The proof follows the spirit of that of Lemma 9. We show that there exists a history where buyer i simply stops buying the item, despite seeing prices below their value.

Lemma 38. *For any PBE of the digital goods setting with discriminatory pricing and any buyer i, assume the first round threshold for i lies on the interior of i's initial support, i.e. $t_i \in (a_i, b_i)$. Then if $\delta > 1/2$, there is some round j and some history $\mathbf{h}^j$ such that:*

1. *$\mathbf{h}^j$ occurs with positive probability in equilibrium, conditioned on $v_i \in [a_i, t_i)$.*

2. for all rounds $1 < j' < j$, buyer i accepts in $\mathbf{h}^j$.

3. all types for i in the support of the updated beliefs F_i^j after $\mathbf{h}^j$ reject in round j.

Proof. We argue by contradiction. For the lemma to be false, it must be that after any history $\mathbf{h}^j$ leading into round j satisfying conditions 1 and 2 in the statement of the Lemma, buyer i has a positive probability of accepting the price p_i^j for round j. Let $T_i^j(v \,|\, \mathbf{h}^j)$ be an indicator variable for whether or not buyer i accepts their round j price assuming the history is $\mathbf{h}^j$ and their value is v. That is, $T_i^j(v \,|\, \mathbf{h}^j) = 1$ if v is less than the threshold t_i^j in that history, and 0 otherwise. By assumption, it must be that

$$\lim_{v \to t_i^-} T_i^j(v \,|\, \mathbf{h}^j) = 1 \tag{15}$$

for any history $\mathbf{h}^j$ satisfying conditions 1 and 2.

For any round $j \geq 2$, we can lower bound buyer i's probability of allocation in round j when they have value v by considering only the contribution from histories satisfying conditions 1 and 2. That is, we may define a lower bound $\underline{X}_i^j(v)$ by:

$$\underline{X}_i^j(v) \equiv \sum_{\mathbf{h}^j \,:\, \mathbf{D}_i^{j'} = A \text{ for all } 1 < j' < j} T_i^j(v \,|\, \mathbf{h}^j) \Pr(\mathbf{h}^j \,|\, v_i = v)$$

Because we have that for any history where buyer i has accepted in every round beyond the second, they accept the next price with positive probability, it must be that as v_i becomes sufficiently high, they accept every price in every on-path history up to round j (excluding the first round). Formally,

$$\lim_{v \to t_i^-} \sum_{\mathbf{h}^j \,:\, \mathbf{D}_i^{j'} = A \text{ for all } 1 < j' < j} \Pr(\mathbf{h}^j \,|\, v_i = v) = 1. \tag{16}$$

Together, facts (15) and (16) imply that $\lim_{v \to t_i^-} \underline{X}_i^j(v) = 1$ for all $j > 1$. Combining these per-round contributions yields a lower bound on the expected total discounted allocation for buyer i. We have that $X_i(v) \geq \sum_{j=2}^{\infty} \delta^{j-1} \underline{X}_i^j(v)$. Note that because $\delta > 1/2$

$$\lim_{v \to t_i^-} \sum_{j=2}^{\infty} \delta^{j-1} \underline{X}_i^j(v) > 1.$$

This implies the existence of a type $v < t_i$ for buyer i such that $X_i(v) > 1$. This is contradicts Lemma 37. □

In a history $\mathbf{h}^j$ satisfying Lemma 38's conditions, buyer i's behavior in round j is such that the seller learns nothing about their type. The no-backtracking condition implies that this buyer will continue to face the same price p_i^j for all subsequent rounds, and will continue rejecting every round. Natural prices implies that there are types for buyer i who will reach this history and will have incentive to deviate: p_i^j will be less than their price, so accepting at least once more will be preferable to rejecting for the rest of the game. We conclude that histories satisfying the conditions of Lemma 38 cannot exist in equilibrium. This, in turn, implies that t_i cannot lie in the support of buyer i's value in the first round, i.e. no learning can occur for buyer i. Since buyer i was chosen arbitrarily, this completes the proof of the Theorem.

G Full Description of Two-Buyer Equilibrium

In this appendix, the two-buyer equilibrium of Section 7.1 in detail. The equilibrium has two phases: an *exploration* phase in which learning occurs, and an *exploitation* phase, where the seller ceases to learn from buying behavior and instead posts the bottom of the strongest player's support for the rest of the game. The exploration phase lasts from the beginning of the game until any buyer rejects a price. The exploitation phase begins when a buyer rejects and continues for the rest of the game.

We first describe the exploration phase of the game. Assume that up to round j, both buyers have accepted all prices. The seller selects a price p_j for the current round, and assumes that the buyers (who have the same distribution F_j after beliefs are updated, supported on $[a_j, b_j]$) will respond according to a threshold t_j satisfying the threshold equation:

$$(t_j - p_j)\left(F_j(t_j) + \tfrac{1-F_j(t_j)}{2}\right) = \frac{t_j - a_j}{2} F_j(t_j)\frac{\delta}{1-\delta}. \tag{17}$$

The lefthand side represents the utility of a buyer i with value t_j when they accept. If the other buyer rejects, which occurs with probability $F_j(t_j)$, then buyer i gets the item with certainty, at price p_j. Otherwise, they get it with probability 1/2 at that same price. The righthand side represents the expected utility from rejection. Rejecting triggers the exploitation phase of the game. If the other buyer rejects as well, then the seller will post a_j for the remainder of the game, and both buyers will accept for the remainder of the game, yielding the righthand side of (17). If the other buyer accepts, the seller will post t_j for the rest of the game, yielding no utility for buyer i.

For a given price, there might exist multiple thresholds satisfying (17). To disambiguate, let $T_{a_j}^{b_j}$ denote the set of solutions to (17). Let p^* be the monopoly price for the initial distribution F. That is, p^* is the value of p maximizing $p(1 - F(p))$.[2] Define the threshold $t_{a_j}^{b_j}(p_j)$ to be p^* if $p^* \in T_{a_j}^{b_j}(p)$, p_j if $p_j \le a_j$, and $t_{a_j}^{b_j}(p) = \max_{t \in T_{a_j}^{b_j}(p_j)} t$ if $T_{a_j}^{b_j}(p_j) \neq \emptyset$ and infinity otherwise.[3] Given this threshold function, the seller optimizes their revenue, which is given by the recurrence:

$$\begin{aligned} R_{a_j}^{b_j}(p_j) = &\left(1 - F_j\left(t_{a_j}^{b_j}(p_j)\right)\right)^2 \left(p_j + \delta R_{t_{a_j}^{b_j}(p_j)}^{b_j}\right) \\ &+ 2F_j\left(t_{a_j}^{b_j}(p_j)\right)\left(1 - F_j\left(t_{a_j}^{b_j}(p_j)\right)\right)\left(p_j + \frac{t_{a_j}^{b_j}(p_j)\delta}{1-\delta}\right) \\ &+ F_j\left(t_{a_j}^{b_j}(p_j)\right)^2\left(\frac{a_j\delta}{1-\delta}\right), \end{aligned}$$

where R_x^y is the seller's optimal revenue from the continuation game where the buyers are both distributed acccording to F_x^y. This equation comes from a straightforward breakdown of the seller's revenue. If both buyers accept, the seller makes revenue p_j, and both buyers must have value at least $t_a^b(p)$, so the seller is faced the next round with a fresh game with values distributed according to F conditioned on $[t_{a_j}^{b_j}(p_j), b_j]$. If exactly one rejects, then the seller receives revenue

[2] If multiple maximizers exist, choose one arbitrarily.

[3] Explicitly seleting monopoly price for a threshold is not required for the strategies we describe to be an equilibrium, but it facilitates the revenue analysis.

p_j and prices at $t_{a_j}^{b_j}(p_j)$ until the end of the game, and this price is accepted. If neither accepts, then the seller prices at a_j for the rest of the game. Note that the seller may solve the above recurrence for the optimal price p_j by value iteration.

We now give a full description of the seller's strategy in Algorithm 10, the buyers' strategy in Algorithm 11, and the threshold selection rule in Algorithm 12. Note that responses to off-path actions will be dictated by careful updates of the beliefs, which we will explain after presenting the strategies.

Algorithm 10: Seller's Strategy

```
Input  : Purchasing history h^j, Support bounds (a_1^j, b_1^j), (a_2^j, b_2^j)
Output: Price p_j
if h^j == (AA)^j then
    a_1^j = a_2^j = a;
    b_1^j = b_2^j = b;
    if a ≥ p* then
        p_j = a
    else
        p_j = arg max_p R_a^b(p)
else
    p_j = max{a_1^j, a_2^j};
```

Algorithm 11: Buyer i Strategy

```
Input  : Purchasing history h^j, Support bounds (a_1^j, b_1^j), (a_2^j, b_2^j), value v_i, price p_j
Output: Purchasing decision for round j
if h^j == (AA)^{j-1} then
    a_j = a_1^j = a_2^j;
    b_j = b_1^j = b_2^j;
    Accept if and only if v_i ≥ t_{a_j}^{b_j}(p_j)
else
    if p_j ≤ max{a_1^j, a_2^j} then
        Accept if and only if v_i ≥ p_j
    else
        Reject
```

We now describe the belief updates. On-path, beliefs are dictated by standard Bayesian updates. In the exploration phase, the buyers have the same support in round j, $[a_j, b_j]$. Any buyers who accept price p_j have their support updated to $[t_{a_j}^{b_j}(p_j), b_j]$. Those who reject have new support $[a_j, t_{a_j}^{b_j}(p_j)]$. In the exploitation phase, each buyer either always accepts or always rejects, so updates are trivial.

Off-path, we use belief updates to implement punishments for deviation. In particular, if all buyers are expected to take the same action but one deviates, the seller will update their beliefs to the maximum value of the support. For the rest of the game, the seller will post this

Algorithm 12: Threshold Selection Rule

Input : Support bounds (a_j, b_j), price p_j

Output: Threshold $t_{a_j}^{b_j}(p_j)$

$p^* = \max_p(1 - F(p))p$;

if $(p^* - p_j)\left(F_j(p^*) + \frac{1-F_j(p^*)}{2}\right) = \frac{p^*-a_j}{2}F_j(p^*)\frac{\delta}{1-\delta}$ **then**

  Return p^*;

else if $p_j \le a_j$ **then**

  Return p_j;

else if $\exists t$ *such that* $(t - p_j)\left(F_j(t) + \frac{1-F_j(t)}{2}\right) = \frac{t-a_j}{2}F_j(t)\frac{\delta}{1-\delta}$ **then**

  Return $\max\{t : (t - p_j)\left(F_j(t) + \frac{1-F_j(t)}{2}\right) = \frac{t-a_j}{2}F_j(t)\frac{\delta}{1-\delta}$;

else

  Return ∞;

value, and with probability 1, all buyers will reject. The details of the belief update algorithm are given in Algorithm 13. Together with Algorithms 10 and 11, this fully specifies equilibrium.

Theorem 39. *The strategies and beliefs specified by Algorithms 10, 11, and 13 is a PBE for* $\delta \ge 2/3$.

In Section G.1, we show that the buyers are best-responding, and in Section G.2, we show that the seller is best-responding. Together, these results prove the theorem.

G.1 Buyer Incentives

As discussed, the belief updates of Algorithm 13 are designed to punish agents for out-of-equilibrium actions. We will use this as a tool for enforcing equilibrium for the buyers. We must first argue that the punishment is effective.

Lemma 40. *If any buyer takes an out of equilibrium action in round* j*, i.e., an action that has zero probability according to the public beliefs, then that buyer does not obtain any utility from rounds* $j + 1$ *and on.*

Proof. As described in Algorithm 13, once a buyer i accepts or rejects when all buyers should have behaved in the opposite fashion according to their strategy, the public belief about this buyer's type becomes a pointmass on the top of their original distribution's support, b. As a result, in all future rounds the seller posts a price of at least b. □

We now argue that the buyers are in equilibrium.

Lemma 41. *For* $\delta \ge 2/3$*, each buyer is best-responding to the strategies of the seller and the other buyer.*

Proof. We break our analysis into two parts: the exploration phase, which occurs when no buyer has rejected yet, and the exploitation phase, where at least one buyer has rejected. These encompass the on-path incentives, but in fact also apply to any history in which only the seller has deviated from equilibrium. If a buyer has taken an off-path action in the history, then Lemma 40 implies that incentives are trivial.

Algorithm 13: Belief updates

Input : Purchasing history h^{j+1}, Current support bounds $(a_1^j, b_1^j), (a_2^j, b_2^j)$, First-round common support bounds (a, b).

Output: Updated support bounds $(a_1^{j+1}, b_1^{j+1}), (a_2^{j+1}, b_2^{j+1})$.

for $i \in \{1, 2\}$ **do**

- **if** *all types for* i *should reject* p_j **then**
  - **if** *buyer* i *rejected* p_j **then**
    - $a_i^{j+1} = a_i^j$, $b_i^{j+1} = b_i^j$
  - **else**
    - $a_i^{j+1} = b$, $b_i^{j+1} = b$
- **else if** *all types should accept* **then**
  - **if** *buyer* i *accepted* **then**
    - $a_i^{j+1} = a_i^j$, $b_i^{j+1} = b_i^j$
  - **else**
    - $a_i^{j+1} = b$, $b_i^{j+1} = b$
- **else**
  - **if** *buyer* i *accepted* **then**
    - $a_i^{j+1} = t_{a_i^j}^{b_i^j}(p_j)$, $b_i^{j+1} = b_i^j$
  - **else**
    - $a_i^{j+1} = a_i^j$, $b_i^{j+1} = t_{a_i^j}^{b_i^j}(p_j)$

Exploration Phase Assume the seller has offered some sequence of prices, all of which have been accepted by both buyers. Then the beliefs about the buyers' values in round j are distributed IID according to F_j, for some support interval $[a_j, b_j]$. Now assume that in the current round, the seller has posted price p_j, and assume that there is at least one threshold t_j satisfying equation (17). We will show that for buyer 1 (without loss of generality), accepting p_j is a best response if $v_1 \geq t_j$, and rejecting is a best response if $v_1 < t_j$.

Assume $v_1 \geq t_j$ The utility for buyer i from accepting, U_A, satisfies

$$u_A \geq \left(F_j(t_j) + \tfrac{1 - F_j(t_j)}{2}\right)(v_1 - p_j) + \frac{\delta}{1 - \delta}(v_1 - t_j)F_j(t_j). \tag{18}$$

The first term comes from the current round: buyer 1 wins at price p with probability 1 if buyer 2 rejects, and with probability 1/2 if buyer 2 accepts. The second term comes from the event that buyer 2 rejects this round, in which case the seller will enter the exploitation phase and offer a price of t_j for the remainder of the game.

Meanwhile, if buyer 1 rejects this round, the seller will enter the exploitation phase regardless of the action of buyer 2. This leaves us with two cases. If buyer 2 rejects, then the seller will post a price of a for the rest of the game (assuming buyer 1 does not attempt further deviations, which by Lemma 40 are not profitable). If buyer 2 accepts, then the seller will instead post t_j in the next round. Note that buyer 1 may now accept t_j next round. Doing so will be profitable, as $v_1 \geq t_j$, but because the seller expects rejection from buyer 1, the seller will update their beliefs for buyer 1 to be a pointmass on the highest possible value, yielding no further utility. It follows that if buyer 1 rejects this round, their expected utility u_R satisfies:

$$u_R \leq F_j(t_j)\frac{v_1 - a_j}{2} \cdot \frac{\delta}{1-\delta} + \delta(1 - F_j(t_j))\frac{v_1 - t_j}{2}. \tag{19}$$

Write $x = v_1 - t_j$. By assumption, $x \geq 0$. By subtracting expression (18) from expression (19) and rearranging, we obtain the following lower bound on the margin by which buyer 1 would prefer to accept:

$$u_A - u_R \geq (t_j + x - p_j)\frac{1 + F_j(t_j)}{2} - (t_j + x - a_j)\frac{F_j(t_j)\delta}{2(1-\delta)} + x\left(\frac{\delta F_j(t_j)}{1-\delta} - \frac{\delta(1 - F_j(t_j))}{2}\right),$$

which we rearrange as:

$$\begin{aligned} &\tfrac{1}{2}\left((F_j(t_j) + 1)(t_j - p_j) - F_j(t_j)(t_j - a_j)\tfrac{\delta}{1-\delta}\right) \\ &+x\left(\tfrac{1+F(t)}{2} - \tfrac{F(t)}{2}\tfrac{\delta}{1-\delta} + \tfrac{\delta F(t)}{1-\delta} - \tfrac{\delta(1-F(t))}{2}\right). \end{aligned} \tag{20}$$

Next, note that we can rearrange the threshold equation (17) to get:

$$(F_j(t_j) + 1)(t_j - p_j) = F_j(t_j)(t_j - a_j)\frac{\delta}{1-\delta}. \tag{21}$$

Applying equation (21) to the expression in (20) allows us to write the utility difference between accepting and rejecting as a product of x and another term which is clearly positive:

$$u_A - u_R \geq x\left(\frac{1 + F_j(t_j)}{2} - \frac{F_j(t_j)}{2}\frac{\delta}{1-\delta} + \frac{\delta F_j(t_j)}{1-\delta} - \frac{\delta(1 - F_j(t_j))}{2}\right). \tag{22}$$

Assume $v_1 < t_j$ If buyer 1 accepts this round, they will win with probability $F_j(t_j) + (1 - F_j(t_j))/2 = (1 + F_j(t_j))/2$. All subsequent prices will be above t_j, so they will receive no continuation payoff. Their utility from accepting is therefore:

$$u_A = \frac{1 + F_j(t_j)}{2}(v_1 - p_j). \tag{23}$$

Meanwhile, if buyer 1 rejects, they trigger the exploitation phase. If buyer 2 also rejects, then the seller offers a for the remainder of the game. If buyer 2 accepts, then the price for the rest of the game is t_j, in which case buyer 1 cannot obtain any utility. The utility from rejecting is therefore:

$$u_R = \frac{F_j(t_j)}{2}\frac{\delta(v_1 - a_j)}{1-\delta}. \tag{24}$$

Combining equations (23) and (24) yields:

$$u_R - u_A = \frac{F_j(t_j)}{2}\frac{\delta(v_1 - a_j)}{1-\delta} - \frac{1 + F_j(t_j)}{2}(v_1 - p_j)$$

After substituting $x = t_j - v_1$ and applying equation (21), we have:

$$u_R - u_A = x\left(\frac{1 + F_j(t_j)}{2} - \frac{F_j(t_j)}{2}\frac{\delta}{1-\delta}\right) \tag{25}$$

Note that we have defined the threshold selection rule such that if $p_j < a_j$, then buyer 1 will accept p_j. Hence, $v_1 < t_j$ only if $p_j > a_j$. In this case, $t_j - p_j < t_j - a_j$, and hence equation (21) implies that the term in parentheses in (25) is positive. It follows that agents with value below t_j will reject.

Assume (17) has no solution It is possible for the price p_j to be such that the threshold equation has no solution. In this case, we require that all buyers reject as a best response.

If buyer 1 accepts, they get utility $u_A = v_1 - p_j$, as buyer 2 will reject this round, and by Lemma 40, they will receive no utility in the future. If they reject, then the seller will enter the exploit phase of the equilibrium and post a_j for the rest of the game. This yields utility $u_R = \frac{\delta}{1-\delta}\frac{v_1 - a_j}{2}$.

When the threshold equation (17) has no solution, it must be that the lefthand side of the rearranged threshold equation (21) is less than the righthand side. This follows from the fact that for $t_j < p_j$, the lefthand side is negative, while the righthand side is positive. Since both sides are continuous functions of t_j, it cannot be that the lefthand side grows larger than the right, or else they would cross to yield a solution. Consequently, if (21) has no solution, it must be that for all t,

$$(1 + F_j(t))(t - p_j) < \frac{\delta}{1-\delta}(t - a_j)F_j(t).$$

Rearranging yields:

$$\begin{aligned}(t - p_j) &< \frac{\delta}{1-\delta}\frac{F_j(t)}{F_j(t)+1}(t - a_j) \\ &\leq \frac{1}{2}\frac{\delta}{1-\delta}(t_j - a).\end{aligned}$$

Setting $t = v_1$ yields that $u_R > u_A$, as desired.

Exploitation Phase: In the exploitation phase, the seller prices at the bottom of the belief support of the strongest buyer. If the seller offers this price or lower, the buyers act as price-takers. If the seller offers a higher price, both buyers reject. We must show that each of these behaviors is a best response.

Seller offers $p \leq \max(a_1, a_2)$ Note that if an agent who is expected to accept chooses instead to reject, they get no utility in the current round, and no utility in the future, by Lemma 40. Accepting the current price is clearly preferable. The only case this does not cover is if $a_1 \neq a_2$ and p lies in between. Assume without loss of generality that $a_1 > a_2$. Buyer 2's purchasing decision does not affect future prices, so they should act as a price taker.

Seller offers $p > \max(a_1, a_2)$ We first argue in the case where $a_1 = a_2 = a$. In this case, the utility of buyer 1 (without loss of generality) from rejecting is $\frac{\delta}{1-\delta}\frac{v_1-a}{2}$, as they will win the item with probability $1/2$ for the rest of the game. If they accept, they will receive utility $v_1 - p$, as by Lemma 40, they will not receive utility in subsequent rounds. Since $\delta \geq 2/3$, we have that $\frac{\delta}{2(1-\delta)} \geq 1$. Since $p > a$ as well, we have that $\frac{\delta}{1-\delta}\frac{v_1-a}{2} > v_1 - p$. Hence, rejection is optimal.

In the case where $a_1 \neq a_2$, the incentive to reject is even greater, due to the fact that the higher-valued agent will receive the item with probability 1 in every subsequent round if they reject. Hence, rejection is optimal here as well.

It follows that in both the exploration and exploitation phases, the buyers are best responding to the strategy of the seller. □

G.2 Seller Incentives

Lemma 42. *The seller is best responding to the actions of the buyers.*

Proof. We break our analysis into three cases: exploration phase, exploitation phase, and off-path analysis.

Exploration Phase In the exploration phase both buyers have the same support, and the seller has three options: price below the common support, post a price p such that there is a threshold response solving equation (17), or post a price p such that no solution to (17) exists. We will show that the second option, posting a price within the common support such that there is a threshold response, is optimal. Since the seller's equilibrium strategy is defined implicitly to be the optimal such price, it will follow that they are best responding in the exploration phase.

Let a_j be the lower bound of the common support. We will first show that pricing below a_j yields less expected revenue than pricing at a_j. This implies that the seller prefers to post a price within the common support. To prove this claim, observe that any price $p_j \leq a_j$ will be accepted with probability 1 by both buyers and cause the beliefs to remain unchanged. Clearly the seller would prefer to induce this outcome with a higher price.

We now argue that posting a price for which a threshold response does not exist is suboptimal. This follows from the fact that both buyers will reject, yielding no revenue and no update to the beliefs. The next round's decision problem is identical to that of the current round, but with payoffs discounted by δ. The seller clearly does not benefit from skipping a round in this manner.

Exploitation Phase In the exploitation phase, the seller posts $\max(a_1^j, a_2^j)$, the higher of the lower bounds of the two buyers supports. We will show that the seller prefers this to their other options, which are posting a lower price, or posting a higher price.

The seller prefers not to post a lower price for the same reason that they would prefer not to post below the common support in the exploration phase. As long as the price p_j satisfies $p_j \leq \max(a_1^j, a_2^j)$, it will be accepted with probability 1, and will not cause the beliefs about the stronger of the two buyers to change. Posting the largest price which induces this outcome is preferable to other such prices.

If the seller posts a price $p_j > \max(a_1^j, a_2^j)$, then this price is rejected by all agents, and beliefs do not change. The seller effectively skips the round and is faced with the same decision next round, with a discount. This is not optimal. Hence, the optimal price to post in the exploitation phase is $\max(a_1^j, a_2^j)$.

Off-Path Analysis We finally analyze the case where a buyer has taken an action in the current history that was not expected of any type. In this case, the seller updates their beliefs to the highest possible value b, and the buyer strategy dictates that they behave as a price-taker. The optimal response to such a scenario is to post b every round. □