Unveiling the Failure of Positive Selection*

Dongkyu Chang† Duk Gyoo Kim‡ Wooyoung Lim§

September 21, 2026

**Abstract**

In dynamic screening problems between an uninformed seller and a privately-informed buyer, theory suggests that the presence of the buyer's outside option leads to a significant surplus for the seller. However, this prediction relies on multiple layers of positive selection reasoning. Therefore, examining whether sellers exhibit positive selection in their belief updates has important managerial implications for the monopoly seller. To investigate this, we conduct a two-round bargaining experiment with finite price alternatives, which enable us to identify first-, second-, and third-order positive selection reasoning. Few subjects adhere to the equilibrium reasoning, and the majority fail even at the first-order reasoning. Consequently, the seller payoffs fall short of the theoretical benchmark.

**Keywords**: Positive Selection, Outside Options, Laboratory Experiments

***JEL* classification numbers**: C78, C91, D03

*We are grateful to Yeon-Koo Che, Ilwoo Hwang, Johannes Hörner, Teddy Kim, and Songfa Zhong for their valuable comments and suggestions. We thank participants at the 2023 KEA Allied Conference and the 16th Maastricht Behavioral and Experimental Economics Symposium, the 5th Annual Workshop of Applied Economics, and the KER International Conference for useful feedback and seminar participants at City University of Hong Kong, Hong Kong Baptist University, Hong Kong University of Science and Technology, Korea Advanced Institute of Science and Technology, Korea University, Seoul National University, Sungkyunkwan University, and Sogang University for their helpful discussions and comments. Chun Yeung Knowlton Chan, Benson Wong, and Ryan Li provided excellent research assistance. This study is supported by a grant from the Research Grants Council of Hong Kong (Grant No. GRF16506019) and the Yonsei University Research Grant of 2025. The Human Participants Research Panel of HKUST has reviewed and approved the safety measure for this study.

†Department of Economics and Finance, City University of Hong Kong. *Email:* `donchang@cityu.edu.hk`

‡School of Economics, Yonsei University. *Email:* `kim.dukgyoo@yonsei.ac.kr`

§Department of Economics, Hong Kong University of Science and Technology. *Email:* `wooyoung@ust.hk`

# 1 Introduction

We investigate a bargaining game involving an uninformed seller and a privately informed buyer, where the buyer's willingness to pay (type) is unknown to the seller. The seller faces the challenge of screening different buyer types over time, a problem known as the dynamic screening problem, without the ability to commit to future prices. Building upon this framework, Board and Pycia (2014)—hereinafter referred to as BP—introduces an outside option for buyers, which could represent an alternative product offered by a third party. Buyers with lower types are more inclined to exercise this outside option and exit the market rather than engage in negotiations with the seller. Consequently, the remaining demand pool primarily consists of higher-type buyers. This *positive selection* emerging among the remaining demand pool motivates the seller to charge a monopoly price that maximizes the expected profit over the prior distribution of buyer types. This pricing strategy expedites the bargaining outcomes, avoiding delays that often undermine the seller's profit in the absence of a commitment to future prices (Coase, 1972). Thus, BP's theoretical analysis reveals that the introduction of the buyer's outside option significantly benefits the seller. Notably, this finding remains robust, as even a minuscule positive value for the buyer's outside option generates qualitatively similar bargaining outcomes.[1]

Positive selection serves as a mechanism that allows a monopoly seller in the market to overcome its lack of commitment power and convert a significant portion of consumer surplus into profit. Consequently, BP's findings have important implications for managerial decision-making as well as market design and regulatory policies across various markets, such as durable-good monopolies, sequential auctions, and lemon markets, which are centered around the dynamic screening problem. If the goal of the market designer is to maintain consumer surplus, BP's results suggest that it is sufficient to prevent buyers from accessing any outside options.[2] Similarly, if the goal of the monopoly's managerial decision-making is to maximize profit, it is essential to create a market environment that provides buyers with outside options. However, this managerial implication appears to contradict the conventional wisdom that restricting monopoly power generally fosters market competition and increases consumer surplus. Therefore, it is crucial to empirically

[1] Catonini (2022) further reinforces BP's findings by proving that the equilibrium strategy is the uniquely strongly rationalizable strategy.

[2] When there is no outside option available, the bargaining game reverts to the standard dynamic screening problem, where the Coase Conjecture (Coase, 1972) holds true. In this case, the uninformed seller does not benefit from inter-temporal price discrimination among different buyer types. The remaining buyers at any given price are more likely to be of lower types, leading to a phenomenon known as *negative selection* in the demand pool. In the absence of commitment power, the seller responds by gradually reducing the offering price over time. Anticipating this price decline, even a high-type buyer tends to delay their purchase, prompting the seller to lower the price even in the early stages to encourage any purchase. As a result, the seller effectively charges the lowest price consistently, earning the minimum possible expected profit in equilibrium. The concept of negative selection has been theoretically explored and confirmed by researchers such as Fudenberg et al. (1985) and Gul et al. (1986). In our companion paper (Chang et al., 2026), we examine the treatment effect of the outside option by comparing environments with and without an outside option.

validate the existence of positive selection before discussing its managerial and policy implications. This justification supports our approach of employing controlled laboratory experiments to obtain empirical evidence.

The primary purpose of our laboratory experiment is to examine whether the experiment participants exhibit positive selection in their belief updates. Our experiment is not merely comparing the predictions of BP with our experimental data: By considering a simple two-round bargaining environment with three discrete types, we can clearly separate out higher-order reasoning from the lower-order ones. Our main interests are to examine the first-, second-, and third-order positive selections. Roughly speaking, the first-order positive selection implies that when a buyer has a low type, he should take the outside option immediately. The second-order positive selection implies that the low- and middle-type buyers do not have an incentive to delay, so the seller believes that in the second round, it is more likely to encounter a high-type buyer. The third-order positive selection implies that the high-type buyer does not have an incentive to delay as well, and the seller, who correctly updates her belief, charges the monopoly price that maximizes the expected earnings of a stage game. We formally derive these predictions of first-, second-, and third-order positive selection by employing the notion of conditional dominance (Shimoji and Watson, 1998). It is a natural consequence that no delay occurs if these orders of rationality hold. Our simple experimental design also allows us to exclude other compounding factors. For example, by restricting the price options to be finite and distinctively different from one another, we enforce the effect of other-regarding preferences to be minimal[3] and the seller's choices are free from ad-hoc tie-breaking rules.

We found that a substantial fraction (51.22%) of the first-round price offers are met with voluntary delay of the negotiation to the second round, while theory predicts that buyers should either accept the offer or exercise the outside option immediately. Examining the seller's posterior belief after the rejection of the first-round offer, we found that only a small proportion of the sellers updated their belief in line with the idea of positive selection: About 37% of the posterior beliefs can be rationalized in the first-order positive selection, about 25% of them can be rationalized in the second-order positive selection, and only about 7% of them can be understood as a result of the third-order positive selection. The steady erosion from 37% to 7% suggests the failure is

[3]For example, in one of the experimental treatments we considered, a seller can offer a price of 10, 180, or 440, where the monopolistic price that maximizes the seller's expected profit is 440. Even if a seller is inequity averse to offering 440 as he feels the matched buyer would earn little, it is too costly to offer the closest price alternative, 180. Also, this alternative offer places the seller in a position of disadvantageous inequity as the offer lets the high-type buyer enjoy more than what the seller could earn. That is, since the prices are distinctively different, the seller's other-regarding preference, even if it exists, cannot play a significant role. From the perspective of the buyer, an outside option always exists. Thus, if the buyer wants to reject the price offer that could yield more payoffs than accepting the outside option merely because the seller's offer does not seem fair, the buyer could have taken the outside option immediately, instead of embracing the risk of random termination and uncertainty about the seller's second(final)-round behavior. Thus, the only sensible case that inequity aversion may alter the buyer's decision is when the buyer's *advantageous* inequity aversion is substantial, which we claim that it is less persuasive. A more discussion is in Section 3.

not localized to one order but accumulates with each additional step of the conditional dominance reasoning. Worse yet, the third-order positive selection predicts a posterior identical to the prior, so a subject who merely retains the prior without reasoning is observationally indistinguishable from one who reasons to the third order. Their second-round price offers resolve the ambiguity: about half of those reporting the equilibrium posterior were naïve players who kept the prior, meaning even the 7% overstates genuine third-order reasoning. As a result, few experiment participants behaved in a way that the BP's logic unfolds. Also, the theory presented in BP and the simplified version of ours predict that the presence of an outside option leads to a substantially large profit for the seller. In our experiment, the seller's average profit turned out to be much lower than they could have attained in theory, which sharply contradicts the main prediction from the positive selection.

The main results invite four natural alternative explanations: (1) that the failure reflects a general inability of strategic reasoning rather than a specific failure of positive selection; (2) that sellers maintain weight on low types because they rationally anticipate buyers' irrational behavior; (3) that other-regarding preferences distort behavior; or (4) that sellers simply lack the competence to identify a profit-maximizing price. We address each through a dedicated supplementary experiment, and none of the four alternatives finds support. In a supplementary experiment *without* outside options (Section 5.1), subjects exhibit negative selection reasoning as the equilibrium calls for it. In a second supplementary experiment where the lowest price offer is unavailable to sellers (Section 5.2), thereby enabling low-type buyers to opt for an outside option immediately, sellers continue to assign weight to the low type. In a third supplementary experiment in which any behavior influenced by other-regarding motives is controlled by the Amended Random Payment method (Lim and Xiong, 2021) (Section 5.3), participants' behavior qualitatively mirrors that of the main experiment. And in a fourth supplementary experiment conducted in a one-shot take-it-or-leave-it setting (Section 5.4), sellers' average offer matches the profit-maximizing price. Section 5 reports these supplementary experiments in detail, and full procedural information appears in Appendix C and the Online Appendix.

The main contribution of this paper is to present and utilize a way to decipher which level of positive selection reasoning fails. This is beyond merely examining the validity of the equilibrium predictions of the asymmetric information bargaining game with an outside option. We found a clear rejection of the equilibrium predictions, and our experimental design allows us to enunciate the drivers of the discrepancies between experimental data and equilibrium predictions. We hope our method of experimentally stripping down the several layers of rational reasoning will be applied in various contexts whose equilibrium predictions are derived from multiple thought processes.

The rest of this paper is organized as follows. In the following subsection, we discuss the closely related literature. Section 2 describes the theoretical environment. Section 3 describes the experimental design and procedure. The main results are reported in Section 4. Section 5 reports four supplementary experiments that rule out alternative explanations. Section 6 concludes.

### 1.1 Literature Review

Positive selection has garnered significant attention in the literature. Board and Pycia (2014) demonstrates that introducing the buyer's outside option, even if its value is arbitrarily close to zero, challenges the Coase conjecture by enabling the seller to earn substantially higher profits than predicted by the conjecture. The role of outside option and positive selection in bargaining is further investigated in Hwang and Li (2017), Chang and Lee (2022), and Fanning (2026), among others.

Similarly, Tirole (2016) shows that a principal can achieve the outcome of a profit-maximizing mechanism without commitment power in a wide range of dynamic screening problems, contradicting the Coase conjecture that implies the principal earns the least profit. These findings highlight the divergent effects of positive selection and negative selection. While negative selection generally harms the principal's interests, positive selection results in the optimal outcome for the principal. We contribute to the literature on the dynamic screening problem by providing experimental evidence.

In the sense that our experimental design allows us to strip down the first-, second-, and third-order positive selection reasoning, we also contribute to the literature on the analysis and identification of higher-order rationality in sequential games. While Kneeland (2015) proposes 'ring games' to identify static depth of reasoning, Ho and Su (2013) extend the level-k framework to dynamic settings to explain failures of backward induction. These approaches measure the *depth* of reasoning a subject can sustain as a general trait. We instead hold the depth requirement fixed and ask which substantive step of one economic mechanism fails: our finite-price design maps each order of positive selection to a specific belief restriction, so a breakdown can be located within the logic of the mechanism rather than attributed to a generic shortfall in reasoning depth.

Many bargaining situations in the real world extend beyond two rounds of offers. Thus, although the two-round experiment in the current paper is best suited to test whether and to what extent subjects exhibit positive selection in their belief updates, it is not proper to test how positive selection (or failure of it) affects the bargaining dynamics and eventual outcomes in the real world based on our experiment. Our companion paper (Chang, Kim and Lim, 2026) addresses this limitation, and the two papers ask distinct questions: the companion estimates the causal effect of the outside option on bargaining dynamics and outcomes in an infinite-horizon setting, whereas the present paper decomposes the belief-updating reasoning behind positive selection and isolates which order of reasoning fails. Neither result can be obtained from the other's design.

## 2 Theoretical Background

### 2.1 Model

The model developed in this section is designed as a benchmark for our laboratory experiment. We deliberately use a finite two-round environment and a finite price set to isolate the positive-selection mechanism of BP rather than to provide a comprehensive description of real-world bargaining. This parsimonious structure makes the reasoning behind positive selection transparent.

**Environment** Consider a two-round version of BP's bargaining game. A seller (he) and a buyer (she) negotiate the price of an indivisible good. The buyer's type (the value of the good) $v$ is private information, which is drawn from a *finite* support $V \subset (0, \infty)$ before the negotiation begins. We often index the buyer types as $v_1 < v_2 < \ldots < v_N$, where $N = |V|$. For any $v' \in V$, let $q(v') = \mathbb{P}\{v = v'\} \in (0, 1)$ denote the probability that the buyer's true type is $v'$. The buyer also has an outside option. The outside option is available in both rounds, and its value is type-independent and given as $w > 0$.[4] Each buyer type's net value is defined as $u(v) := v - w > 0$, assumed to be strictly positive. The seller's valuation of the good is normalized as zero; the seller has no outside option.

In the first round $t = 1$, the seller first makes a price offer $p_1$ from the *finite* set of feasible offers $\mathcal{P} \subset \mathbb{R}_+$; the assumption imposed on $\mathcal{P}$ will be discussed shortly. Then, the buyer chooses to accept this offer, exercise the outside option, or reject both the offer and the outside option to delay the final decision to the next round. In the second round $t = 2$, if it occurs, the seller again makes a price offer $p_2 \in \mathcal{P}$, and then the buyer chooses to accept this offer, exercise the outside option, or reject both the offer and the outside option.

In addition to the buyer's and the seller's decisions as described above, Nature may step in with probability $\epsilon \in [0, 1)$ at the end of the first round. In the case that Nature intervenes, it overrides any decision by the buyer in the first round and forces the negotiation to be delayed to the second round.[5] When the second round occurs, the seller cannot distinguish whether it was the buyer's voluntary decision or Nature's intervention that brought the negotiation to the second round; the seller updates his posterior beliefs, being aware of both possibilities.[6]

[4] We assume that outside options are type-independent for the sake of simplicity. BP's original model also allows the outside option to be fully type-dependent, including the case with a type-independent outside option as a special case.

[5] Note that we consider both cases $\epsilon = 0$ and $\epsilon > 0$. Taking the possibility of Nature's intervention into account, if the buyer chooses to accept $p_1$ or exercise the outside option in the first round, the bargaining still can be delayed to the second round with probability $\epsilon$. The bargaining is delayed to the second round for sure if the buyer chooses to delay in the first round.

[6] When $\epsilon > 0$, the seller's second-round information set is reached with positive probability conditional on any first-round offer $p_1$. Given the prior $q$ at the start of the $p_1$-continuation game, Bayes' rule therefore determines the posterior for any given buyer strategy. This feature will help us obtain unique, testable predictions for the seller's posterior beliefs both on and off the equilibrium path.

The game can end in the first round only if Nature does not intervene after the buyer's response. In that case, the game ends if the buyer accepts $p_1$ or exercises the outside option. If Nature intervenes, or if the buyer *voluntarily* delays, the game proceeds to the second round. The second round is final: the buyer either accepts $p_2$, exercises the outside option, or rejects both.[7] At any terminal history in which the buyer accepts $p_t$ or exercises the outside option in round $t \in \{1, 2\}$, the buyer's and seller's payoffs are, respectively,

$$\begin{aligned} \delta^{t-1}(v - p_t) &\quad \text{and} \quad \delta^{t-1} p_t && \qquad \text{if the buyer accepts } p_t, \\ \delta^{t-1} w &\quad \text{and} \quad 0 && \qquad \text{if the buyer exercises the outside option,} \end{aligned}$$

where $\delta \in (0, 1)$ is the common discount factor. If the buyer rejects both $p_2$ and the outside option in round 2, both players receive zero payoff.

**Assumptions** We impose several assumptions on the set of feasible prices $\mathcal{P}$ and the seller's payoff function. First, we assume

$$\begin{aligned} &(i) \quad \mathcal{P} \cap \{u(v) : v \in V\} = \varnothing; \\ &(ii) \quad \mathcal{P} \cap (u(v), u(v')) \neq \varnothing \quad \text{for any } v, v' \in V \text{ such that } v < v'. \end{aligned} \tag{A1}$$

The first condition in (A1) requires that no feasible price makes any buyer type exactly indifferent between accepting the offer and exercising the outside option. We impose this assumption to avoid indeterminacy from the buyer's tie-breaking rule. The second condition requires that $\mathcal{P}$ contains a sufficiently fine grid, so that, for any two buyer types $v < v'$, we can find $p \in \mathcal{P}$ such that $u(v) < p < u(v')$.

To state the second assumption, define

$$\overline{p}(v) := \max\{p \in \mathcal{P} : u(v) = v - w > p\}$$

as the highest feasible price that buyer type $v \in V$ strictly prefers to the outside option. We assume that the set $\{p \in \mathcal{P} : u(v) = v - w > p\}$ is nonempty, so that $\overline{p}(v)$ is well-defined for every $v \in V$.[8] Note that (A1)-($ii$) implies that $\overline{p}(v)$ strictly increases in $v$. We assume that the largest gap between

[7] Rejecting both is strictly inferior because $w > 0$, so the buyer's second-round choice is effectively between accepting $p_2$ and exercising the outside option. Assumption (A1)-($i$) below rules out indifference between the two, so each buyer type has a unique best response in the second round.

[8] Assumption (A1)($ii$) guarantees that $\{p \in \mathcal{P} : u(v) > p\}$ is nonempty for every buyer type except possibly the lowest type $v_1$. If this set were empty for $v_1$, exercising the outside option would strictly dominate both accepting and voluntarily delaying after every $p_1$. Hence, type $v_1$ would never trade and would not affect the seller's optimal price choices. We assume nonemptiness for $v_1$ only to avoid treating this trivial case separately.

$u(v)$ and $\overline{p}(v)$ across buyer types is sufficiently small:

$$\Delta := \max_{v \in V}\{u(v) - \overline{p}(v)\} < \frac{1-\delta}{\delta} w. \tag{A2}$$

This assumption holds, for example, if the feasible prices are closely spaced over the relevant range.

Lastly, let the seller's *static monopoly payoff* from price $p \in \mathcal{P}$ be

$$\Pi(p) := p \sum_{v \in V : u(v) > p} q(v). \tag{2.1}$$

We impose the following single-peakedness condition. The static payoff $\Pi$ has a unique maximizer $p^* \in \mathcal{P}$. Moreover, $\Pi$ is strictly increasing on feasible prices below $p^*$ and strictly decreasing on feasible prices above $p^*$. That is, for any $p', p'' \in \mathcal{P}$,

$$\begin{aligned} p' < p'' \leq p^* \quad &\Longrightarrow \quad \Pi(p') < \Pi(p''), \\ p^* \leq p' < p'' \quad &\Longrightarrow \quad \Pi(p') > \Pi(p''). \end{aligned} \tag{A3}$$

Define

$$v^* := \max\{v \in V : \overline{p}(v) \leq p^*\}.$$

This is the highest buyer type whose highest acceptable feasible price is weakly below the static monopoly price $p^*$. Note that every buyer type accepts any feasible price weakly below $\overline{p}(\min V)$, so that $\Pi(p) = p$ is strictly increasing on $\{p \in \mathcal{P} : p \leq \overline{p}(\min V)\}$; hence $p^* \geq \overline{p}(\min V)$. As a result, $\{v \in V : \overline{p}(v) \leq p^*\}$ is nonempty, and therefore $v^*$ is well-defined.

**Strategies and Posterior Beliefs** For each buyer type $v \in V$, let $\sigma_v(A|p_1)$ denote the probability that type $v$ accepts $p_1$ in the first round, conditional on the seller offering $p_1$. Similarly, let $\sigma_v(O|p_1)$ and $\sigma_v(D|p_1)$ denote the probabilities that type $v$ exercises the outside option and voluntarily delays, respectively. The buyer's second-round decision is straightforward. For any $p_2 \in \mathcal{P}$, type $v$ accepts $p_2$ if $u(v) > p_2$ and exercises the outside option otherwise. We may therefore omit the buyer's second-round behavioral strategy without loss when describing her equilibrium strategy. Let $\sigma_B = \langle \sigma_v : \{A, O, D\} \times \mathcal{P} \to [0,1] \rangle_{v \in V}$ generically denote the buyer's behavioral strategy.

For any $p_1$ and $p_2$, $\sigma_S(p_1)$ denotes the probability that the seller offers $p_1$ in the first round, and $\sigma_S(p_2|p_1)$ denotes the probability of offering $p_2$ in the second round, conditional on the negotiation reaching round 2 after the first-round offer $p_1$. We denote by

$$\hat{q}(\tilde{v}|p_1) = \mathbb{P}\{v = \tilde{v} \mid \text{the negotiation reaches round 2 after } p_1\}$$

the seller's posterior belief that the buyer type is $\tilde{v}$ after the negotiation reaches the second round following the first-round offer $p_1$. We write $\hat{q} = \{\hat{q}(\cdot|p_1)\}_{p_1 \in \mathcal{P}}$ for the seller's posterior belief system.

**The Role of Perturbation $\epsilon > 0$** Before proceeding to the analysis, we clarify why Nature's intervention is included. Our goal is not only to derive the equilibrium outcome, but also to test the seller's posterior beliefs after continuation. In the unperturbed game with $\epsilon = 0$, as we will prove shortly, the positive-selection logic implies that no buyer type voluntarily delays. Without perturbation, therefore, Bayes' rule alone does not determine the seller's posterior after continuation. This is not an artifact of our two-round formulation. In BP's original infinite-horizon model, positive selection likewise implies that no buyer type delays in equilibrium, so continuation is off path there as well.

In our perturbed game with $\epsilon > 0$, this issue does not arise conditional on a given first-round price. After any first-round price, every buyer type can reach the second round because Nature may force continuation. Hence, for any given strategy of the buyer, the seller's posterior is pinned down by Bayes' rule. Restrictions on the buyer's strategy, whether from equilibrium or from strategic dominance reasoning, translate into testable restrictions on the seller's posterior.

## 2.2 Conditional Dominance

We use conditional dominance, which evaluates strict dominance conditional on reaching each information set (Shimoji and Watson, 1998), to derive implications for the buyer's decision to delay and, through Bayes' rule, for the seller's posterior belief after continuation[9]. The basic idea is to identify buyer types whose voluntary delay cannot survive iterated conditional dominance. When $\epsilon > 0$, once voluntary delay is eliminated for a given type, its probability of reaching round 2 is exactly $\epsilon$, whereas the continuation probability of a type whose delay has not been eliminated may exceed $\epsilon$. These differences in continuation probabilities generate positive selection in the seller's posterior belief.

To formulate this argument, we apply iterated conditional dominance separately to each continuation game after the seller's initial offer. Fix any first-round price $p_1 \in \mathcal{P}$ and consider the continuation game beginning immediately after $p_1$ has been offered. At the beginning of this continuation game, the seller's belief about the buyer's type is the prior $q$.[10] In this game, which we call the $p_1$-*continuation game*, $p_1$ is treated as exogenously given, and the game starts with the buyer's response to $p_1$. All other elements of the original game, including Nature's possible intervention, remain unchanged. If the negotiation reaches round 2, the seller chooses a second-round price $p_2$, after which the buyer makes her second-round decision.

We analyze these continuation games rather than applying conditional dominance to the entire

[9] Shimoji and Watson (1998) focus on extensive-form games without moves of Nature. Appendix A adapts their framework to our setting by representing buyer types as separate buyer-agents and incorporating the type draw and Nature's intervention into expected payoffs.

[10] Because the seller has no information about the buyer's type when choosing $p_1$, conditioning on his own price choice does not change this belief. This is consistent with the no-signaling-what-you-don't-know condition of Fudenberg and Tirole (1991).

game because our object is to understand the buyer's behavior and the seller's belief conditional on each realized first-round price, not about the rationalizability of the seller's initial price choice. In a deletion procedure applied to the entire game, some first-round prices may themselves be eliminated for the seller, after which the buyer's response to those prices need not be evaluated. Our experiment, however, elicits posterior beliefs after realized first-round prices, including prices that need not survive such a procedure. Indeed, conditional on a realized first-round price $p_1$, the $p_1$-continuation game is precisely the continuation of the laboratory game that we analyze. Treating $p_1$ as exogenously given therefore means conditioning on the realized history, not assuming that the experiment exogenously assigns the seller's first-round price.

Appendix A provides the formal argument; here, we sketch the basic logic. Before analyzing voluntary delay, we first eliminate suboptimal buyer responses in round 2. At each second-round information set, it is uniquely optimal for buyer type $v$ to accept $p_2$ if $u(v) > p_2$ and to exercise the outside option otherwise. Hence, at that information set, any strategy prescribing another response is strictly dominated by an otherwise identical strategy prescribing the unique optimal response. After these preliminary deletions, the only buyer decision relevant to the subsequent analysis is her response $A$, $O$, or $D$ to $p_1$.

Index the buyer types as $v_1 < \cdots < v_N$. The argument proceeds recursively from the lowest type to the highest: eliminating voluntary delay for type $v_k$ uses only the prior elimination for types below $v_k$. Consider first the lowest type $v_1$. In the second round, every price $p_2 < \bar{p}(v_1)$ is strictly dominated by $\bar{p}(v_1)$ at the seller's second-round information set. Indeed, $u(v_k) \geq u(v_1) > \bar{p}(v_1) > p_2$ for every type $v_k$, so every buyer type accepts both prices, while $\bar{p}(v_1)$ gives the seller strictly more revenue.[11] Once such prices have been eliminated, for any remaining second-round price $p_2$, let $c(v_1, p_2) := \delta \max\{w, v_1 - p_2\}$ denote type $v_1$'s continuation payoff. Since $p_2 \geq \bar{p}(v_1)$,

$$c(v_1, p_2) \leq C(v_1) := \delta \max\{w, v_1 - \bar{p}(v_1)\} = \delta[v_1 - \bar{p}(v_1)].$$

The last equality follows from $\bar{p}(v_1) < u(v_1) = v_1 - w$. Against the same second-round price $p_2$, choosing voluntary delay yields type $v_1$ the payoff $c(v_1, p_2)$, whereas exercising the outside option after $p_1$ yields $(1-\epsilon)w + \epsilon c(v_1, p_2)$. Hence, the payoff gain from exercising the outside option rather than voluntarily delaying is $(1-\epsilon)[w - c(v_1, p_2)]$, which is at least $(1-\epsilon)[w - C(v_1)]$. Moreover, from Assumption (A2),

$$C(v_1) = \delta[w + u(v_1) - \bar{p}(v_1)] \leq \delta(w + \Delta) < w.$$

Therefore, exercising the outside option strictly dominates voluntary delay for type $v_1$.

The same logic applies recursively, though the argument differs according to whether $\epsilon = 0$ or

[11] If $\bar{p}(v_1)$ is already the lowest feasible price, this seller-side deletion is vacuous. Our experiment is designed to have this property.

$\epsilon > 0$. Consider first $\epsilon = 0$, and suppose voluntary delay has been eliminated for types $v_1, \ldots, v_{k-1}$. These lower types can then no longer reach round 2. Hence, at any surviving strategy profile that reaches the seller's second-round information set, some type $v_j \geq v_k$ must have voluntarily delayed. Every such type accepts $\bar{p}(v_k)$, so any second-round price $p_2 < \bar{p}(v_k)$ is strictly dominated by $\bar{p}(v_k)$ at the seller's second-round information set. Once these lower prices have been eliminated, the payoff comparison used for $v_1$ implies that exercising the outside option strictly dominates voluntary delay for type $v_k$. Iterating eliminates voluntary delay for every buyer type.

When $\epsilon > 0$, eliminating voluntary delay for lower types does not exclude them from round 2, because Nature may still force continuation. To compare a second-round price $p_2 < \bar{p}(v_k)$ with $\bar{p}(v_k)$, distinguish the two ways in which round 2 can be reached: either Nature intervenes, which occurs with probability $\epsilon$ for every type, or Nature does not intervene and the buyer voluntarily delays, which can occur only for types $v_j \geq v_k$. Conditional on Nature's intervention, the seller's discounted expected revenue from offering $p_2$ is $\delta\Pi(p_2)$; hence, if $p_2 < \bar{p}(v_k) \leq p^*$, Assumption (A3) implies that $\bar{p}(v_k)$ yields strictly greater expected revenue than $p_2$. If Nature does not intervene but the buyer voluntarily delays, every buyer type that may have delayed is at least $v_k$ and therefore accepts both $p_2$ and $\bar{p}(v_k)$, so replacing $p_2$ with $\bar{p}(v_k)$ also increases the seller's revenue whenever this event occurs. Thus, every second-round price below $\bar{p}(v_k)$ is strictly dominated by $\bar{p}(v_k)$ at the seller's second-round information set, and the payoff comparison used for $v_1$ then eliminates type $v_k$'s voluntary delay.

This argument requires $\bar{p}(v_k) \leq p^*$, so it applies only up to the cutoff type $v^*$. Beyond that cutoff, Assumption (A3) no longer guarantees that raising the price to $\bar{p}(v_k)$ increases the seller's expected revenue conditional on Nature's intervention.

Motivated by this recursive structure, fix $p_1 \in \mathcal{P}$ and define the *low-to-high deletion chain for voluntary delay* as follows. The first step removes second-round prices below $\bar{p}(v_1)$ and then eliminates type $v_1$'s voluntary delay. Recursively, given that voluntary delay has been eliminated for types $v_1, \ldots, v_{k-1}$, the *k-th step* removes any remaining second-round prices below $\bar{p}(v_k)$ and then eliminates type $v_k$'s voluntary delay. We refer to the outcome of the first $k$ steps, that is, the elimination of voluntary delay for types $v_1, \ldots, v_k$, as the *k-th-order deletion of voluntary delay*, or simply the *k-th-order deletion*. Let

$$K^* := \max\{k \in \{1, \ldots, N\} : \bar{p}(v_k) \leq p^*\},$$

so that $v_{K^*} = v^*$.

**Proposition 1** (Deletion of voluntary delay)**.** *Suppose Assumptions* (A1)*,* (A2) *and* (A3) *hold.*

*(i) Suppose $\epsilon = 0$. In every $p_1$-continuation game, the low-to-high deletion chain yields the $k$-th-order deletion of voluntary delay for every $k = 1, \ldots, N$.*

*(ii) Suppose $\epsilon > 0$. In every $p_1$-continuation game, the low-to-high deletion chain yields the $k$-th-order deletion of voluntary delay for every $k = 1, \ldots, K^*$. In particular, if $K^* = N$, voluntary delay is eliminated for every buyer type after every first-round price.*

*Proof.* See the Appendix. □

The low-to-high chain in Proposition 1 transparently displays the logic of positive selection, but it is not the only way to organize the deletions. Shimoji and Watson (1998) instead define a maximal procedure that removes all currently conditionally dominated strategies at each step and therefore does not presuppose any particular deletion order.[12] Appendix A.3 shows that the maximal procedure yields the same conclusions as Proposition 1: when $\epsilon = 0$, it eliminates voluntary delay for every buyer type; when $\epsilon > 0$, it eliminates voluntary delay for every buyer type weakly below $v^*$.

Importantly, the first step of the deletion argument is particularly robust. Once the relevant round-2 actions have been deleted, voluntary delay by the lowest buyer type is conditionally dominated regardless of which delay strategies remain for the higher types. Thus, the elimination of the lowest type's delay is not an artifact of imposing a low-to-high deletion order; it also follows from the order-independent maximal procedure. Consequently, when $\epsilon > 0$, the corresponding first-order posterior-belief restriction, derived in the next subsection, is likewise not an artifact of the imposed deletion order.

### 2.3 Posterior-Belief Implications

The low-to-high deletion chain suggests that observing continuation should shift relative likelihood away from lower buyer types and toward higher buyer types. We first formalize this directional property, which we call $k$-th-order positive selection, and then show that the $k$-th-order deletion of voluntary delay implies it.

For any posterior belief system $\hat{q} = \{\hat{q}(\cdot|p_1)\}_{p_1 \in \mathcal{P}}$, define the posterior-to-prior ratio

$$R_i(p_1) := \frac{\hat{q}(v_i|p_1)}{q(v_i)} \qquad \forall i = 1, \ldots, N, \quad \forall p_1 \in \mathcal{P}.$$

We say that $\hat{q}$ satisfies *$k$-th-order positive selection* if, for every $p_1 \in \mathcal{P}$,

$$R_1(p_1) \leq \cdots \leq R_k(p_1) \leq 1, \tag{2.2}$$

$$R_k(p_1) \leq R_j(p_1) \qquad \forall j = k+1, \ldots, N. \tag{2.3}$$

[12] More generally, Chen and Micali (2013) show that different deletion orders may leave different sets of surviving strategies but induce the same terminal histories in their setting.

The ratio $R_i(p_1)$ measures how observing continuation reweights type $v_i$ relative to the prior. We use posterior-to-prior ratios because the prior need not be uniform. Comparing posterior probabilities directly would mix two effects: differences in prior probabilities and the information conveyed by continuation. The posterior-to-prior ratio isolates the latter. Condition (2.2) says that the first $k$ types are weakly downweighted relative to the prior and that the degree of downweighting becomes weakly larger as the buyer's type decreases. Condition (2.3) says that every type above $v_k$ is reweighted at least as strongly as $v_k$. Because $\epsilon > 0$ and the prior has full support, every Bayes-consistent posterior in the perturbed game has full support. Hence, for any $i \leq k$ and $j > i$, the inequality $R_i(p_1) \leq R_j(p_1)$ is equivalent to

$$\frac{\hat{q}(v_j|p_1)}{\hat{q}(v_i|p_1)} \geq \frac{q(v_j)}{q(v_i)}.$$

Thus, observing continuation weakly raises the posterior odds of every higher type relative to each of the first $k$ types. Together, the two conditions (2.2) and (2.3) say that continuation shifts relative likelihood away from lower types and toward higher types.

The inequalities (2.2) imply that positive selection is nested: if $\hat{q}$ satisfies $k$-th-order positive selection, it also satisfies $k'$-th-order positive selection for every $k' \leq k$. At the highest order, these inequalities imply that the posterior coincides with the prior. Indeed, setting $k = N$ gives $R_i(p_1) \leq 1$, or equivalently $\hat{q}(v_i|p_1) \leq q(v_i)$, for every $i = 1, \ldots, N$. Since the prior and posterior probabilities both sum to one, we necessarily have $\hat{q}(\cdot|p_1) = q$ in this case.

We now show that the low-to-high deletion chain, combined with Bayes' rule, generates this form of positive selection. For this purpose, we focus on $\epsilon > 0$. When $\epsilon = 0$, Proposition 1 eliminates voluntary delay for every buyer type, so round 2 is off path and Bayes' rule places no restriction on the seller's posterior.

Fix any buyer behavioral-strategy profile $\sigma_B = \langle \sigma_{v_\ell} \rangle_{\ell=1}^{N}$. Conditional on the buyer's type being $v_j$, the negotiation reaches round 2 after $p_1$ with probability $\epsilon + (1-\epsilon)\sigma_{v_j}(D|p_1)$. Bayes' rule gives the posterior associated with $\sigma_B$ by weighting these continuation probabilities by the prior and normalizing across types. We denote this Bayesian posterior by $\hat{q}^{\sigma_B}$:

$$\hat{q}^{\sigma_B}(v_j|p_1) := \frac{q(v_j)\left[\epsilon + (1-\epsilon)\sigma_{v_j}(D|p_1)\right]}{\epsilon + (1-\epsilon)\sum_{\ell=1}^{N} q(v_\ell)\sigma_{v_\ell}(D|p_1)} \qquad \forall j = 1, \ldots, N, \quad \forall p_1 \in \mathcal{P}. \tag{2.4}$$

Suppose the low-to-high deletion chain yields the $k$-th-order deletion of voluntary delay in every $p_1$-continuation game. Any buyer behavioral-strategy profile surviving this deletion must satisfy

$$\sigma_{v_i}(D|p_1) = 0 \qquad \forall i = 1, \ldots, k, \quad \forall p_1 \in \mathcal{P}.$$

Substituting these restrictions into (2.4) yields

$$\hat{q}^{\sigma_B}(v_i|p_1) = \frac{q(v_i)\epsilon}{\epsilon + (1-\epsilon)\sum_{\ell=k+1}^{N} q(v_\ell)\sigma_{v_\ell}(D|p_1)} \qquad \forall i = 1,\dots,k, \quad \forall p_1 \in \mathcal{P}. \tag{2.5}$$

We say that a posterior belief system $\hat{q}$ is *compatible with the $k$-th-order deletion of voluntary delay* if $\hat{q} = \hat{q}^{\sigma_B}$ for some buyer behavioral-strategy profile $\sigma_B$ surviving that deletion.

**Proposition 2** (Posterior-belief implication)**.** *Suppose $\epsilon > 0$. Every posterior belief system compatible with the $k$-th-order deletion of voluntary delay satisfies $k$-th-order positive selection. Moreover, for every $p_1 \in \mathcal{P}$,*

$$R_1(p_1) = \cdots = R_k(p_1). \tag{2.6}$$

*Proof.* See the Appendix. □

Equation (2.6) is sharper than the inequality-based definition of positive selection. Because Nature's intervention is type-independent, all types whose voluntary delay has been eliminated have the same posterior-to-prior ratio. This equality is not part of the definition: it is an additional implication of the particular perturbation used in our model.

Neither the definition nor the proposition imposes any ordering among types above $v_k$. Their posterior-to-prior ratios depend on their respective probabilities of voluntary delay. Moreover, the downweighting of the first $k$ types need not be strict. If no higher type voluntarily delays, continuation occurs only through Nature's intervention and the posterior coincides with the prior. If at least one higher type voluntarily delays with positive probability, every type whose delay has been eliminated is strictly downweighted relative to the prior.

## 2.4 Equilibrium

We now turn from the reasoning behind positive selection to the equilibrium outcome of the two-round game. The central prediction of BP is that the buyer's outside option generates positive selection and thereby allows the seller to charge the static monopoly price $p^*$ in an essentially unique equilibrium, even without commitment. The next proposition shows that this prediction holds in the unperturbed game with $\epsilon = 0$ and is robust to sufficiently small positive perturbations. We use perfect Bayesian equilibrium (hereafter, PBE or equilibrium), as defined in Fudenberg and Tirole (1991, Definition 8.2).

**Proposition 3.** *Suppose Assumptions (A1)–(A3) hold. Then the following statements hold.*

*(i) If $\epsilon = 0$, every PBE induces the following first-round behavior after any offer $p_1 \in \mathcal{P}$: each buyer type $v \in V$ accepts $p_1$ if $u(v) > p_1$ and exercises the outside option otherwise. In particular, no buyer type voluntarily delays. The seller offers $p^*$ in the first round.*

*(ii) For every sequence $\epsilon^n \downarrow 0$, every associated sequence of PBE, and every $p_1 \in \mathcal{P}$,*

$$\lim_{n\to\infty} \sigma_v^n(A|p_1) = 1 \qquad \forall v \in V \text{ such that } u(v) > p_1,$$
$$\lim_{n\to\infty} \sigma_v^n(O|p_1) = 1 \qquad \forall v \in V \text{ such that } u(v) < p_1,$$

*where $\sigma_v^n(\cdot|p_1)$ denotes type $v$'s first-round behavioral strategy in the PBE of the game with perturbation $\epsilon^n$. Consequently, $\lim_{n\to\infty} \sigma_v^n(D|p_1) = 0$ for every $v \in V$, so the induced first-round behavior converges to the $\epsilon = 0$ outcome in part (i).*

*(iii) There exists $\bar{\epsilon} > 0$ such that, for every $\epsilon \in [0, \bar{\epsilon})$, the seller offers $p^*$ in the first round with probability one in every PBE.*

*Proof.* See the Appendix. □

To summarize the section, we have presented three results. Proposition 1 identifies the conditional dominance deletions underlying positive selection, and Proposition 2 translates these deletions into posterior-belief restrictions. Proposition 3 then shows that BP's equilibrium prediction survives in our environment and is robust to a small positive perturbation. Together, these results confirm that our finite two-round environment reproduces both BP's equilibrium prediction and the positive-selection reasoning that supports it. The simplifications we impose therefore do not alter BP's central insight, so this environment provides a setting in which their mechanism can be tested. Beyond these results, conditional on any first-round price $p_1$ being offered, a positive perturbation ensures that the seller's second-round information set is reached with positive probability, yielding testable restrictions on his posterior beliefs. This feature is absent from BP's original framework, in which continuation is off path in equilibrium.

# 3 Experimental Design and Hypotheses

## 3.1 Experimental Design

As considered in Section 2, we consider a two-round ($t = 1, 2$) bargaining game between a seller and a buyer with one-sided private information in our experiment. The primary purpose of this experimental design is to examine whether the participants in the experiment exhibit positive selection in their belief updates. For that purpose, we elicit the sellers' posterior beliefs whenever the negotiation reaches the second round.

The seller has an indivisible good for sale. The buyer's value for the good, $v$, is drawn before the negotiation from the finite support $V = \{v_L, v_M, v_H\}$ and is privately observed by the buyer. We refer to $v_L$, $v_M$, and $v_H$ as the low, middle, and high types, respectively. The buyer's outside option is $w$, which is commonly known and independent of the realization of $v$. We use the parameter values[13]

$$v_L = 70, \quad v_M \in \{90, 240, 420\}, \quad v_H = 500, \quad \text{and} \quad w = 50.$$

The prior is uniform:

$$q(v_L) = q(v_M) = q(v_H) = \frac{1}{3}.$$

In each round $t = 1, 2$, the seller offers a price from

$$\mathcal{P} = \{v_L - w - \Delta,\ v_M - w - \Delta,\ v_H - w - \Delta\},$$

where we set $\Delta = 10$ to avoid payoff ties. For notational simplicity, define

$$p_k := v_k - w - \Delta \qquad \forall k \in \{L, M, H\},$$

so that $\mathcal{P} = \{p_L, p_M, p_H\}$. Under this price set, $p_k$ is the highest feasible price at which type $v_k$ strictly prefers acceptance to the outside option. Thus, in the notation of Section 2, $p_k = \overline{p}(v_k)$ for all $k = L, M, H$. By construction, accepting $p_k$ yields strictly more than the outside option for type $v_k$ and all higher types, but strictly less for all lower types.

Given this price set, the remaining actions and payoffs follow the model described in Section 2. To implement discounting and make delay costly, we use the random-termination method of Roth and Murnighan (1978), with continuation probability $\delta = 0.8$.[14] Finally, we set $\epsilon = 0.001$ so that Nature's intervention has a negligible effect on subjects' incentives while allowing us to obtain testable posterior-belief predictions after continuation.

When the negotiation reaches the second round,[15] the seller is asked to report her belief about the matched buyer's type and offer the second-round price. When asking the sellers to report their beliefs, we emphasized that the reported probabilities are not shared with the buyer, and the sole purpose is to help them think about an appropriate price offer in the second round.

[13] As explained below, the experimental design includes three treatments, corresponding to the three possible values of $v_M$. The values $v_L = 70$ and $v_H = 500$ remain constant across treatments.

[14] The qualitative theoretical predictions are unchanged over a broad range of values of $\delta$, although $\delta$ must be strictly below one. See Section 3.2.

[15] We did not adopt the strategy method, which asks all the seller subjects to report what they would have believed if the first-round offer were rejected. Although it has the advantage of collecting posterior belief data from all seller subjects, we argue that the disadvantages of the strategy method outweigh the advantages for two reasons. First, it is unnatural to ask the experiment participants to report their *posterior* belief *prior* to the event. The experiment instructions would have been unnecessarily complicated if the posterior belief had been asked. Second, the strategy method may create an unwanted experimenter demand effect. The main purpose of our experiment is to examine the higher-order positive selection reasoning, and enforcing the seller subjects to think about the posterior belief in the case of the first-round price offer being rejected could confound the interpretations of experimental observations. Since a meta-analysis study (Brandts and Charness, 2011) shows mixed results on the treatment effect comparison between the strategy method and direct-response method, while a treatment effect found with the strategy method is also observed with the direct-response method, we took the conservative approach.

Table 1: Experimental Design

| *M90* | *M240* | *M420* |
|---|---|---|
| $v \in \{70, 90, 500\}$ | $v \in \{70, 240, 500\}$ | $v \in \{70, 420, 500\}$ |

* Each participant has ten newly paired matches.
* Continuation probability to the next round is 0.8.
* Buyer's value $v$ is uniformly drawn from $\{70, v_M, 500\}$.

It is worth noting that we deliberately did not incentivize the elicitation of subjects' beliefs for several reasons. There are known issues with complicated belief-elicitation methods that do not reliably elicit truthful belief reports (Burdea and Woon, 2022; Danz et al., 2022). Additionally, while there may be some noise in unincentivized belief reports, arguably there are no substantial reasons to believe subjects would systematically misreport their beliefs. By emphasizing that the purpose of belief reporting is to help subjects make appropriate decisions in the second round, we believe our elicitation is simple yet reasonable enough to induce subjects' cognitive effort.[16]

Our experimental design is summarized in Table 1. Three treatment conditions differ in the value of the middle type, so we call *M90* if the middle type's value is 90. *M240* and *M420* are called accordingly. Our experiment was conducted using oTree (Chen et al., 2016) at HKUST. A total of 294 subjects[17] were recruited from the graduate and undergraduate population of the university. We had 5 sessions per treatment. Each session consisted of 16 to 22 participants, and we had 106, 100, and 88 participants in treatments *M90*, *M240*, and *M420*, respectively. We used the random matching protocol and between-subject design. The participants' roles were fixed across matches. In all sessions, subjects participated in one practice match[18] and ten payment-relevant matches of the bargaining game under one treatment condition. Participants were reshuffled to form new pairs after each match so that there are no strategic dynamics between matches. One of the ten payment-relevant matches was randomly selected for payment. The average payment is HKD 139 ($\approx$ USD 17.8) including the HKD 40 show-up fee. Sample experimental instructions can be found in Appendix D.

[16] If we view this belief reporting as communication between one's current and future selves, then truthful reporting becomes more compelling (Burchardi and Penczynski, 2014). Given the indifference between lying and truthful reporting, a lexicographic preference for honesty (Demichelis and Weibull, 2008) would lead to truthful belief reports. The preference for truth-telling is well-documented in the literature (Gneezy et al., 2018; Abeler et al., 2019).

[17] This counts only those who participated in the main experiment. Additional 86, 90, 44, and 72 subjects participated in the supplementary experiments reported in Sections 5.1, 5.2, 5.3, and 5.4, respectively.

[18] We exclude this practice match from our analysis, as we clarified that the purpose of the practice match is to familiarize themselves with the experiment, and their performance in that practice round is not relevant for payment. Including this practice match data does not change any results in a meaningful way.

### 3.2 Rationale for the Choice of Experimental Parameters

In choosing $\Delta$ and other parameter values, we make sure the low buyer type finds that delaying the negotiation in the first round is strictly dominated by taking the outside option:

$$\delta[v_L - (u(v_L) - \Delta)] = \delta(w + \Delta) < (1-\epsilon)w + \delta\epsilon w \quad \Longleftrightarrow \quad \delta\Delta < (1-\epsilon)(1-\delta)w. \tag{3.1}$$

We impose this condition to facilitate the first step of the conditional-dominance logic described in Section 2, which, through Bayes' rule, gives rise to first-order positive selection in the seller's posterior belief. From equation (3.1), we restrict $\delta$ to be strictly smaller than 1, and $\Delta$ to be sufficiently small. Given $w = 50$, $\delta = 0.8$, and $\epsilon \approx 0$, $\Delta < 12.5$. So we set $\Delta = 10$. Note that our choice of parameters here is also consistent with the assumption (A2) in Section 2.

The preceding payoff comparison is especially transparent in our experiment because of the sparse price set. Compared to the general model discussed in Section 2, our experiment restricts $\mathcal{P}$ to only three prices, each of which is $\Delta$ below one buyer type's net value, and there is no feasible price below $p_L$. Thus, the first seller-price deletion in the general conditional-dominance argument is vacuous in our experiment: even if the low type expects the most favorable feasible second-round price $p_L$, delaying yields at most $\delta(w + \Delta)$, which is already strictly lower than taking the outside option immediately. This makes the first step of the conditional-dominance logic particularly clear.

The first-step payoff comparison is also robust to the buyer's aversion to *disadvantageous inequity* (Fehr and Schmidt, 1999). The lowest-type buyer compares the payoff from taking the outside option in the first round with the payoff from delaying in the hope of accepting $p_L$ in the second round. Our choice of parameter values guarantees that the lowest-type buyer is in an advantageous position in both outcomes.[19] Thus, any aversion to disadvantageous inequity cannot overturn the first-order positive-selection logic.[20]

Also, given these parameter values, we make sure that offering a price that is too low should not be a profit-maximizing strategy for the seller. Specifically, we require the following two conditions to hold:

$$\Pi(p_L) = p_L < \Pi(p_H) = p_H \cdot q(v_H), \tag{3.2}$$

$$\Pi(p_L) = p_L < \Pi(p_M) = p_M \cdot [q(v_H) + q(v_M)]. \tag{3.3}$$

[19] The buyer's exercise of the outside option yields payoffs of $w = 50$ for the buyer and 0 for the seller. If the buyer accepts $p_L$, the buyer and the seller, respectively, obtain $v_L - p_L$ and $p_L$, where $(v_L - p_L) - p_L = v_L - 2(v_L - w - \Delta) = 50 > 0$.

[20] The remaining concern is that advantageous inequity aversion could make delay attractive because delay creates a chance of random termination that yields both players equal zero payoffs. We regard this channel as less persuasive because it requires a sufficiently strong aversion to advantageous payoff differences and may be attenuated by risk aversion, efficiency-seeking preferences, or other behavioral biases. Some previous studies also report that inequity aversion does not significantly affect bargaining dynamics (Bereby-Meyer and Niederle, 2005; Kim, 2023). The ARP design in Section 5.3 provides a direct test by eliminating other-regarding motives from the payoff-relevant decision environment.

where $\Pi$ denotes the seller's static monopoly profit defined in (2.1). If neither (3.2) nor (3.3) holds, offering $p_L$ yields the seller at least as much static profit as either targeted higher price. Since $p_L$ is acceptable to all buyer types, the low-price pooling outcome would then become the relevant benchmark: all buyers accept the offer immediately, leaving no room for screening, voluntary delay, or nontrivial posterior-belief updating. Both conditions hold strictly under our parameter values, so sellers face a genuine tradeoff between selling broadly at a low price and targeting higher types at a higher price.

Given this tradeoff, evaluating the static monopoly price $p^*$ under our experimental parameters yields

$$p^* = \begin{cases} p_H = 440 & \text{if } v_M = 90 \text{ or } 240, \\ p_M = 360 & \text{if } v_M = 420. \end{cases}$$

By Proposition 3, this static monopoly price is the equilibrium first-round offer for sufficiently small $\epsilon$; thus, the equilibrium first-round price is $p_H = 440$ in *M90* and *M240*, but $p_M = 360$ in *M420*. We vary the value of the middle type so that the equilibrium price offer varies accordingly. The same distinction also affects the posterior-belief implications of positive-selection reasoning. When $p^* = p_H$, the static monopoly price is the highest feasible price, so the positive-selection logic rules out voluntary delay by all three buyer types. When $p^* = p_M$, the conditional-dominance logic directly rules out voluntary delay for the low and middle types, while the high type requires a separate continuation argument after the off-path high offer $p_H$. We discuss these posterior-belief implications in the next subsection.

### 3.3 Theoretical Predictions and Hypotheses

Table 2 summarizes the equilibrium predictions for each treatment. In all treatments, the seller offers the static monopoly price identified in the previous subsection, and buyers either accept the offer or take the outside option; no voluntary delay should occur. The predictions are identical in $M90$ and $M240$, while $M420$ yields a different equilibrium price. Thus, in addition to comparing $M240$ and $M420$, we also test for a null treatment effect between $M90$ and $M240$, which provides an additional check on whether observed treatment effects are consistent with the theory.

Note that Table 2 provides the outcome-level equilibrium predictions. Since these equilibrium predictions are built upon multiple layers of positive-selection reasoning, we dissect them to examine which parts of the reasoning are supported by the experimental evidence. In particular, Section 2 shows that conditional-dominance deletions imply posterior-belief restrictions through Bayes' rule; the hypotheses below translate those restrictions into the three-type environment of our experiment. In what follows, to simplify the notation, we write $\sigma_H$, $\sigma_M$, and $\sigma_L$ as shorthand for $\sigma_{v_H}$, $\sigma_{v_M}$, and $\sigma_{v_L}$, respectively.

| | $M90$ | $M240$ | $M420$ |
|---|---|---|---|
| Seller offers† | 440 | 440 | 360 |
| Buyer with $v_L$ plays† | O | O | O |
| Buyer with $v_M$ plays† | O | O | A |
| Buyer with $v_H$ plays† | A | A | A |
| Expected Payoffs of Seller | 146.67 | 146.67 | 240 |
| (Ex ante) Expected Payoffs of Buyer | 53.33 | 53.33 | 83.33 |

O: takes the outside option. A: accepts the offer

†: If the negotiation reaches round 2, the same prediction applies.

Table 2: Summary of Equilibrium Predictions

**First-Order Positive Selection** The "minimal" rationality we require is that the low type should never choose voluntary delay. As discussed in Section 3.2, our parameter choice makes delaying in the first round strictly dominated for the low type by taking the outside option, even when the low type expects the most favorable feasible second-round price. Hence, the first step of the positive-selection reasoning leads to

$$\sigma_L(D|p_1) = 0 \qquad \forall p_1 \in \mathcal{P}.$$

Given $\sigma_L(D|p_1) = 0$, a seller who applies this minimal positive-selection logic should attribute the presence of a low-type buyer in the second round only to Nature's intervention. Middle- and high-type buyers, by contrast, may reach the second round either through Nature's intervention or through voluntary delay. Therefore, by Bayes' rule, the posterior probability assigned to the low type should be weakly below its prior. Since the prior is uniform in our experiment, this also implies that the posterior belief assigned to the low type should be weakly below the posterior belief assigned to either the middle or high type.

**Hypothesis 1** (First-order positive selection)**.** *No low-type buyers choose voluntary delay ($\sigma_L(D|p_1) = 0$), and thus $\hat{q}(v_L|p_1) \leq q(v_L) = 1/3$ for any $p_1 \in \mathcal{P}$. If the game proceeds to the second round, the posterior belief assigned to each higher type is weakly greater than the posterior belief assigned to the low type, for any price offer in round 1. That is, $\hat{q}(v_L|p_1) \leq \min\{\hat{q}(v_M|p_1), \hat{q}(v_H|p_1)\}$ for any $p_1 \in \mathcal{P}$.*

As discussed in Section 3.2, under our parameter values the no-delay component of Hypothesis 1 is robust to disadvantageous inequity aversion. A remaining concern is that advantageous inequity aversion could make delay attractive; we address this concern through the ARP supplementary experiment in Section 5.3.

**Second-Order Positive Selection** The next step of positive-selection reasoning applies the same logic to the middle type. Given Hypothesis 1, a seller who applies the first-order positive-selection logic will not find it optimal to offer $p_2 = p_L$ in the second round. Intuitively, once the low type is not expected to voluntarily delay, the price $p_L$ no longer attracts enough additional demand compared to $p_M$, as guaranteed by (3.3). Therefore, a middle-type buyer who delays cannot rationally expect a second-round price below $p_M$. Then, even under the most favorable second-round price $p_M$, delaying yields at most $\delta(v_M - p_M) = \delta(w + \Delta)$, which is strictly lower than taking the outside option immediately by (3.1). Thus, the middle type also should not choose voluntary delay:

$$\sigma_M(D|p_1) = 0 \qquad \forall p_1 \in \mathcal{P}.$$

Together with Hypothesis 1, this means that low- and middle-type buyers do not voluntarily reach the second round; their presence in the second round is attributed only to Nature's intervention. Any posterior beliefs that are consistent with this reasoning therefore satisfy

$$\hat{q}(v_L|p_1) = \hat{q}(v_M|p_1) \leq \frac{1}{3} \leq \hat{q}(v_H|p_1).$$

This leads us to our second hypothesis.

**Hypothesis 2** (Second-order positive selection)**.** *No low- or middle-type buyers choose voluntary delay ($\sigma_M(D|p_1) = \sigma_L(D|p_1) = 0$). Thus, $\hat{q}(v_M|p_1) \leq q(v_M) = 1/3$ and $\hat{q}(v_L|p_1) \leq q(v_L) = 1/3$ for any $p_1 \in \mathcal{P}$. If the game proceeds to the second round, the posterior belief assigned to the high type is weakly greater than the posterior belief assigned to either lower type, for any price offer in round 1. That is, $\hat{q}(v_L|p_1) = \hat{q}(v_M|p_1) \leq \hat{q}(v_H|p_1)$ for any $p_1 \in \mathcal{P}$.*

**Third-order Positive Selection** The final step of positive-selection reasoning concerns the high type. In our three-type environment, third-order positive selection corresponds to full positive selection: no buyer type chooses voluntary delay. If this is the case, then the negotiation reaches the second round solely because of Nature's intervention. Since Nature's intervention is independent of the buyer's type, the seller's posterior belief coincides with the prior:

$$\hat{q}(v_L|p_1) = \hat{q}(v_M|p_1) = \hat{q}(v_H|p_1) = \frac{1}{3}.$$

First, consider the cases $v_M = 90$ and $v_M = 240$. In both cases, the static monopoly price is the highest feasible price ($p^* = p_H = 440$). Given the first- and second-order positive-selection reasoning, the seller's posterior assigns weakly more weight to the high type than the prior, so $p_H$ remains strictly optimal for the seller in the second round. Hence, a high-type buyer cannot benefit from voluntary delay: after $p_1 < p_H$, accepting immediately is better, and after $p_1 = p_H$, delay only discounts the same price. Thus, the high type also should not choose voluntary delay.

**Hypothesis 3.A** (Third-order positive selection)**.** *Suppose $v_M \in \{90, 240\}$, so that $p^* = p_H$. In addition to the no-delay predictions for the low and middle types in Hypotheses 1 and 2, no high-type buyers choose voluntary delay: $\sigma_H(D|p_1) = 0$ for all $p_1 \in \mathcal{P}$. If the negotiation reaches the second round, the posterior belief equals the prior for all types and any first-round price. In equilibrium, the seller offers $p^* = p_H$ in round 1, the high-type buyer accepts $p^*$, and the low- and middle-type buyers exercise the outside option.*

Next, consider the case $v_M = 420$. In equilibrium, the seller offers $p_1 = p^* = p_M = 360$. In what follows, we discuss what happens both on and off the path. First, suppose that the seller offers $p_1 \in \{p_L, p_M\}$ in the first round. Given the first- and second-order positive-selection reasoning, the seller's second-round offer is not expected to fall below $p_1$. Hence, a high-type buyer cannot benefit from voluntary delay: after $p_1 = p_L$, accepting immediately is better than waiting, and after $p_1 = p_M$, delay only leads to the same or a higher price. Thus, for $p_1 \in \{p_L, p_M\}$, no buyer type should choose voluntary delay. If the negotiation reaches the second round, the posterior belief should coincide with the prior.

Now, suppose that the seller offers $p_1 = p_H$. In this case, the high type's incentive to delay depends on the seller's second-round offer. If the high type expects $p_2 = p_H$, accepting $p_H$ immediately is better than delaying. If, however, the high type always accepts $p_H$, continuation is attributed to Nature. Then, the seller's posterior coincides with the prior, and the seller offers the static monopoly price $p_M$ in the second round. Anticipating this, the high type would prefer to delay. Conversely, if the high type always delays, the seller infers that the buyer is very likely to be the high type and offers $p_H$, making delay unattractive. Hence, in the continuation equilibrium, the high type mixes after $p_1 = p_H$, and the seller mixes between $p_M$ and $p_H$ in the second round. The two mixing probabilities are pinned down by their indifference conditions. Under our parameter values, the high type chooses voluntary delay with probability

$$\sigma_H(D|p_1 = p_H) \approx 0.0035.$$

As a result, the seller's equilibrium posterior belief after continuation is

$$\hat{q}(v_L|p_1 = p_H) = \hat{q}(v_M|p_1 = p_H) = \frac{2}{13}, \qquad \hat{q}(v_H|p_1 = p_H) = \frac{9}{13}.$$

This posterior belief does not satisfy the third-order positive-selection restriction defined in Section 2.3. This is consistent with Proposition 2: we have $p^* = p_M < p_H$ when $v_M = 420$, so the conditional-dominance logic directly rules out voluntary delay only for the low and middle types, but not for the high type.

**Hypothesis 3.B** (Third-order positive selection)**.** *Suppose $v_M = 420$ so that $p^* = p_M$. (i) If $p_1 \in \{p_L, p_M\}$, then no high-type buyers choose voluntary delay ($\sigma_H(D|p_1) = 0$). Together with*

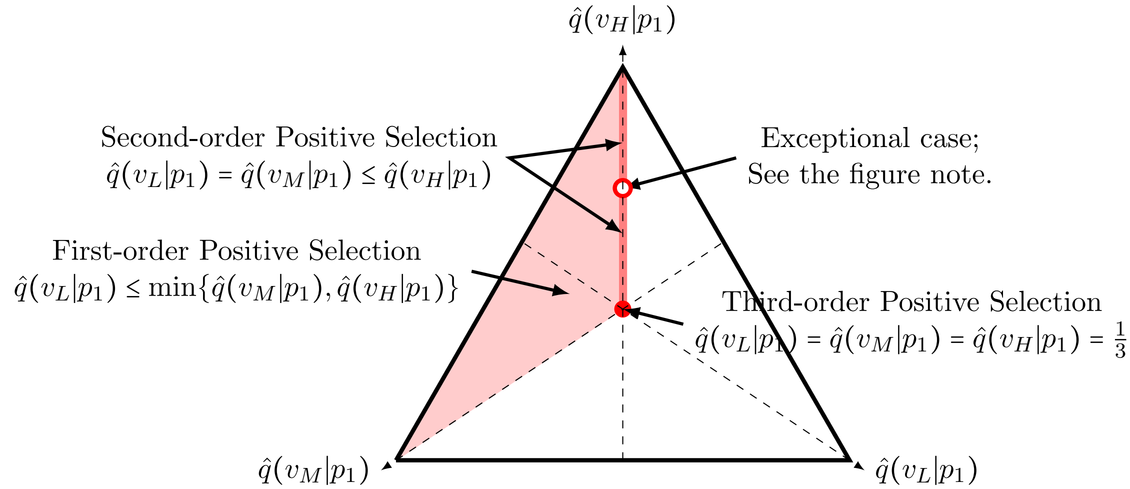


Figure 1: Positive Selections and Posterior Beliefs

The unit simplex represents all possible posterior beliefs after the first-round rejection. Each dashed line indicates the axis for the corresponding belief. The shaded area and the vertical line respectively represent the set of posterior beliefs admissible under the first- and second-order positive selection. The filled circle at the center represents the posterior belief under the third-order positive selection, where the posterior coincides with the prior. The hollow circle near the top vertex represents the posterior belief predicted after $p_1 = p_H$ in treatment *M420*.

*Hypotheses 1 and 2, this means that no buyer type chooses voluntary delay. If the negotiation reaches the second round, the posterior belief is the same as the prior. (ii) If $p_1 = p_H$, then high-type buyers choose voluntary delay with a small probability ($\sigma_H(D|p_1 = p_H) \approx 0.0035$). If the negotiation reaches the second round after $p_1 = p_H$, the posterior belief that a buyer is a high type is nearly 70%:*

$$\hat{q}(v_L|p_H) = \hat{q}(v_M|p_H) = \frac{2}{13}, \qquad \hat{q}(v_H|p_H) = \frac{9}{13}.$$

*In equilibrium, the seller offers $p_1 = p^* = p_M$, and thus no voluntary delay occurs on the equilibrium path.*

Figure 1 illustrates the ranges of posterior beliefs associated with first-, second-, and third-order positive selection reasoning. Any posterior belief $(\hat{q}(v_L|p_1), \hat{q}(v_M|p_1), \hat{q}(v_H|p_1))$ can be represented as a point on the unit simplex. The first-order positive selection implies that $\hat{q}(v_L|p_1)$ is smaller than both $\hat{q}(v_M|p_1)$ and $\hat{q}(v_H|p_1)$, which is represented by the top-left shaded triangle. The second-order positive selection restricts posterior beliefs to $\hat{q}(v_L|p_1) = \hat{q}(v_M|p_1) \leq \hat{q}(v_H|p_1)$, represented by the vertical line from the top vertex to the center of the unit simplex. Third-order positive selection restricts posterior beliefs to the center point, $\hat{q}(v_L|p_1) = \hat{q}(v_M|p_1) = \hat{q}(v_H|p_1)$. The hollow circle near the top vertex represents the posterior belief after $p_1 = p_H$ in *M420*, where the seller assigns probability 9/13 to the high type and 2/13 to each lower type.

It is worth noting that these hypotheses are nested. First-order positive selection restricts the seller's posterior belief about the low type; second-order positive selection adds the corresponding restriction for the middle type; and third-order positive selection, in our three-type environment, implies that no buyer type chooses voluntary delay and hence that the posterior coincides with the prior. Thus, except for the case $p_1 = p_H$ in *M420*, satisfying a higher-order hypothesis also implies satisfying the lower-order ones. This nested structure allows us to identify which posterior-belief implication of positive-selection reasoning fails when the experimental data depart from the equilibrium benchmark.

**The Seller's Price Offers** Suppose that the negotiation reaches the second round after the first-round offer $p_1$. Given any posterior beliefs $\hat{q}(\cdot|p_1)$, a seller who maximizes expected payoff in the second round chooses $p_2$ to solve

$$\max_{p_2 \in \mathcal{P}} \sum_{v : u(v) \geq p_2} p_2 \cdot \hat{q}(v|p_1) \quad \forall p_1 \in \mathcal{P}. \tag{3.4}$$

Since we ask sellers to report their posterior beliefs before choosing $p_2$, we can test whether their second-round prices are optimal given their reported beliefs.

**Hypothesis 4.** *Given the seller's reported posterior belief $\hat{q}(\cdot|p_1)$, the price offered in the second round maximizes the seller's expected payoff in (3.4).*

By testing Hypotheses 1–4 together, we can assess not only whether the equilibrium prediction is supported but also where positive-selection reasoning may break down. If, for example, Hypotheses 1–3.B are supported but Hypothesis 4 is rejected, then it is the seller's failure to choose the payoff-maximizing price in the second round given his self-reported posterior beliefs. On the other hand, if Hypothesis 4 is supported but any of Hypotheses 2, 3.A, and 3.B are rejected, then it is the seller's failure of higher-order positive-selection reasoning. If Hypothesis 1 is rejected, then the failure occurs already at the first step of positive-selection reasoning.

So far, the hypotheses above describe how sellers would update their beliefs and choose prices. The next hypothesis concerns the buyers' decisions. In theory, if the buyer expects that $p_2$ will not be lower than $p_1$, she should accept $p_1$ whenever $v - w > ep_1$ and take the outside option otherwise. In the following, we set our hypothesis based on two important factors that may affect buyers' decisions: misspecified beliefs and inequity aversion.

For all three treatment conditions *M90*, *M240*, and *M420*, theory predicts that $\sigma_v(D|p_1)$ is exactly or very close to zero for any $(v, p_1) \in V \times \mathcal{P}$.[21] A buyer would find such no-delay strategies optimal only if

$$\mathbb{E}[\delta \max\{v - p_2, w\}] \leq \max\{w, v - p_1\}, \tag{3.5}$$

[21] In the case of $v_M \in \{240, 90\}$, we have $\sigma_v(D|p_1) = 0$ for any $(v, p_1) \in V \times \mathcal{P}$. In the case of $v_M = 420$, we still have $\sigma_v(D|p_1) = 0$ for all $(v, p_1)$ except for $(v_H, p_H)$, in which case $\sigma_{v_H}(D|p_1 = p_H)$ also remains very close to 0.

where the expectation on the left-hand side is taken with respect to the buyer's expectation about the seller's second-round offer $p_2$. If this inequality fails, that is, if some buyers are mistakenly optimistic about $p_2$, then they may find voluntary delay optimal, contrary to our theoretical predictions.

Table 3 shows the cases in which (3.5) can be violated. When either $v = v_L$ or $p_1 = p_L$, the inequality always holds regardless of the buyer's misspecified beliefs. In the remaining four cases (shaded in red or blue), the inequality can be violated if the buyer believes that $p_2$ is highly likely to be low.[22] Since the violation condition for $p_M$ (shaded in blue) is more stringent than that for $p_H$ (shaded in red), if the experimental data shows some "delays," then it is most likely in the $v$–$p_1$ pair shaded in red, somewhat likely in the pair shaded in blue, and not likely in other pairs. This leads to the following hypothesis.

| $\mathbb{E}[\delta \max\{v - p_2, w\}] \le \max\{w, v - p_1\}$? | | | |
|---|---|---|---|
| $v \backslash p_1$ | $p_L$ | $p_M$ | $p_H$ |
| $v_L$ | always holds | always holds | always holds |
| $v_M$ | always holds | can be violated | can be violated |
| $v_H$ | always holds | can be violated | can be violated |

Table 3: Validity of not expecting "too low price" in round 2

**Hypothesis 5.** *Buyers with $v_M$ or $v_H$ are more likely to choose voluntary delay in response to $p_H$, somewhat likely to delay in response to $p_M$, and not likely to delay in response to $p_L$. Buyers with $v_L$ are not likely to delay in response to any price offer in round 1.*

Testing Hypothesis 5 enables us to gather insights into whether the buyer's decision to delay is influenced by their misspecified expectation of the second-round price offer. If the buyer's expectation violates (3.5), it would imply that the buyer anticipates the second-round price offer to be sufficiently low, justifying the decision to delay. If our observations are consistent with Hypothesis 5, it would indicate that buyers' decisions are rational given their misspecified beliefs. Conversely, a lack of consistency might suggest that their decisions are driven by some other fundamental reasoning failure.

Finally, in equilibrium, the seller's first-round offer must be the static monopoly price $p^*$. No voluntary delay occurs in equilibrium.

[22] For example, (3.5) is violated, and thus a buyer of type $v = v_M$ would reject $p_1 = p_M$ iff $\Delta b_M + (v_M - p_L - w)b_L > [w(1-\delta) + \Delta]/\delta$, where $b_M$ (respectively, $b_L$) represents the buyer's (misspecified) probability that $p_2 = p_M$ (respectively, $p_2 = p_L$). Similarly, the buyer would reject $p_1$ in the cases corresponding to the other three colored cells iff a weighted sum of $b_L$ and $b_M$ is sufficiently large. The exception occurs in the treatment condition M90 for the case $(v, p) = (v_H, p_M)$, in which (3.5) holds for any belief of the buyer, and thus a rejection should not occur. The intuition is as follows: with $v_M = 90$, the price $p_1 = p_M = 30$ is already attractive to the high-type buyer. Therefore, the high-type buyer never finds rejecting $p_1 = p_M$ optimal, regardless of her beliefs about $p_2$.

**Hypothesis 6.** *The observed first-round offer is the static monopoly price $p^*$. No voluntary delay occurs.*

## 4 Results

In this section, we first report the summary of experimental findings and then present each part of experimental findings in the corresponding order of our hypotheses. Throughout this section, we say that a buyer *voluntarily delays* the first-round seller price $p_1$ when she declines both the offer and the outside option, thereby continuing to the second round through her own choice. Unless it causes confusion, we use 'reject' interchangeably with voluntary delay; in particular, 'rejecting $p_1$' never refers to exercising the outside option, which we always call exit or taking the outside option.

| | $M90$ | $M240$ | $M420$ |
|---|---|---|---|
| Avg.Offer (Theory) | 346.25 (440) | 295.86 (440) | 378.50 (360) |
| %Delay_$v_L$ (Theory) | 39 (0) | 35 (0) | 32 (0) |
| %Delay_$v_M$ (Theory) | 51 (0) | 60 (0) | 74 (0) |
| %Delay_$v_H$ (Theory) | 63 (0) | 50 (0) | 59 (0) |
| Avg.SellerPayoffs (Theory) | 54.43 (146.67) | 87.94 (146.67) | 165.16 (240) |
| Avg.BuyerPayoffs (Theory) | 111.60 (53.33) | 119.20 (53.33) | 86.25 (83.33) |

Table 4: Summary of Experimental Findings

Notes: Avg.Offer is the average of the first-round price offers. %Delay_$v_i$ is the percentage that a buyer with type $v_i$ voluntarily delays the price offer. Avg.SellerPayoffs and Avg.BuyerPayoffs are the average payoffs of the sellers and the buyers, respectively. Theoretically predicted values are in parentheses. The number of observations for $M90$, $M240$, and $M420$ are 530, 500, and 440, respectively. Given the stark contrast between the theoretical predictions and the experimental data, statistical tests (binomial tests for rejection proportions and $t$-tests for price offers) confirm that all observed averages, with the exception of the buyer's payoff in $M420$, are significantly different from the theoretical benchmarks at the $p < 0.01$ level.

Table 4 shows the overall experimental findings, juxtaposing theoretical predictions. Except for a few coincidental numbers, there are substantial differences between experimental findings and theoretical predictions. The seller's equilibrium price offer is the smallest in $M420$, but the seller's average price offer shows the opposite. The seller's equilibrium price offer in $M90$ is identical to that in $M240$, but the average offer in $M90$ is significantly larger than that in $M240$. The average payoff[23] of the seller is the largest in $M420$, but that is significantly smaller than the expected payoff. The buyer's average payoff in $M420$ was the smallest, while theory predicts the opposite.

A substantial proportion of voluntary delays is one of the most distinctive inconsistencies we observe from the data. For any buyer type and any price in the first round, the unique equilibrium

[23]Since one of the ten matches was randomly selected for payment, the average payoff (the average of ten matches) is different from what the experiment participants actually earned.

predicts that every buyer should either accept the first-round price offer or exercise the outside option, so no one should choose to delay. Even when a buyer draws the lowest value so that rejecting does not create a chance for any practically profitable deviations, 39% of the low-type buyers voluntarily delay in $M90$, 35% and 32% of the low-type buyers delay in $M240$ and $M420$, respectively. Middle- and high-type buyers tend to delay more than 50% of the price offer in the first round.

Since we collect the seller's posterior belief about the matched buyer's type when the first round ends up with the buyer's voluntary delay, we could exploit the luxury of 614 observations about the posterior beliefs.[24] Figure 2 shows all reported posterior beliefs. A larger circle indicates more observations in the center of the circle.

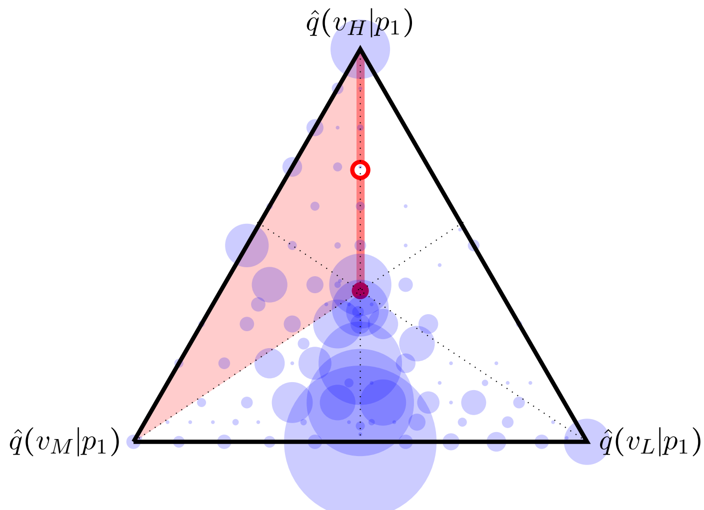


Figure 2: Observed posterior beliefs

The most frequent posterior belief is at $(\hat{q}(v_L|p_1), \hat{q}(v_M|p_1), \hat{q}(v_H|p_1)) = (0.5, 0.5, 0)$. The second- and third-most frequent posterior beliefs are $(0.45, 0.45, 0.1)$ and $(0.4, 0.4, 0.2)$, and these three beliefs account for 19% (128 out of 614) of the entire observed posterior beliefs. In Figure 2, the posterior beliefs spotted in the top-left shaded area of the unit simplex are the ones that can be rationalized in the first-order positive selection, and about two-thirds of posterior belief reports are outside of the area. This evidence already indicates that the majority of sellers fail at the first-order positive selection reasoning. Note further that the posterior being in the shaded area is a necessary, but not a sufficient, condition supporting the first-order positive selection, so we interpret our observations most favorably toward equilibrium reasoning.

**Result 1.** *51.22% (753 out of 1,470) of the first-round price offers were met with voluntary delay. Among 614 reported posterior beliefs, 36.64% of them are rationalized in the first-order positive*

[24] The bargaining procedure moves on to the second round with probability $\delta = 0.8$, so the 614 posterior beliefs are about 81.5% of the bargaining cases in which the seller's first-round offer is met with voluntary delay.

*selection.*

It is worth noting that the majority of posterior beliefs seem to reflect the *negative selection*, believing that the high-type buyers were more likely to accept the first-round price offer so the remaining demand pool consists of more of the middle- and low-type buyers. This reasoning theoretically holds when the outside option does not exist, and it leads to a qualitatively opposite prediction about the equilibrium price dynamics and the seller's payoff.Also, if the sellers believed that the high-type buyers would accept the price offer and the low-type buyers would take the outside option in the first round, that is, if they applied the positive and the negative selection reasoning simultaneously, then they would have reported that the remaining demand pool mainly consists of the middle type. As illustrated in Figure 2, few posterior beliefs are reported around the vertex of $\hat{q}(v_M|p_1)$, indicating that the positive selection reasoning had not been applied. Thus, we claim the sellers' failure of positive selection reasoning[25] is the main driving force behind the stark discrepancy between the theoretical predictions and our experimental findings.

One might worry that this pattern reflects not a failure of positive selection specifically but a more general inability in belief-updating reasoning required after a first-round rejection. We address this concern in Section 5.1, which reports a supplementary experiment in which the buyer has no outside option, while the seller has the same action set. In that environment, the equilibrium reasoning calls for negative selection[26] and we find that subjects update their beliefs in a manner broadly consistent with the equilibrium prediction. The failure documented in our main experiment is therefore specific to the higher-order reasoning that positive selection requires.

We further investigate how many reported posterior beliefs satisfy the second- and third-order positive-selection predictions. The second-order deletion implies the sharper restriction $\hat{q}(v_L|p_1) = \hat{q}(v_M|p_1) \leq \hat{q}(v_H|p_1)$. We use a weaker criterion and count the posterior-belief reports that satisfy $\hat{q}(v_L|p_1) \leq \hat{q}(v_H|p_1)$ and $\hat{q}(v_M|p_1) \leq \hat{q}(v_H|p_1)$.[27]

**Result 2.** *24.59% of the posterior beliefs (151 out of 614) are rationalized in the second-order positive selection.*

Except for the case of $p_1 = p_H$ in $M420$ where the posterior belief under the third-order positive selection is $(2/13, 2/13, 9/13)$, the equilibrium posterior belief of all other cases is $(1/3, 1/3, 1/3)$. Since we asked the seller subjects to report how likely their matched buyer is of high/middle/low type on an integer scale of 0 to 100, we did not take the prior belief $(33.\dot{3}, 33.\dot{3}, 33.\dot{3})$ as a

[25]Some buyers' behaviors were not rational as well, but even if the sellers would take into account some buyers' irrational behaviors, the reported posterior beliefs should have been different from what we observe. See the discussions after Result 4.

[26]In the no-outside-option environment, standard Coasean bargaining dynamics apply. See Appendix B for formal derivation.

[27]This criterion is weaker not only than the sharp restriction (2.6) in Proposition 2, but also than the inequality-based definition of second-order positive selection in (2.2) and (2.3). Under the uniform prior, the latter is equivalent to $\hat{q}(v_L|p_1) \leq \hat{q}(v_M|p_1) \leq 1/3 \leq \hat{q}(v_H|p_1)$.

valid answer. Subjects who wanted to submit an equal likelihood in the posterior belief should then submit a permutation of $(33, 33, 34)$ or similar. To count the incidences of the third-order positive selection, we regard any posterior beliefs such that $\max\{\hat{q}(v_L|p_1), \hat{q}(v_M|p_1), \hat{q}(v_H|p_1)\} - \min\{\hat{q}(v_L|p_1), \hat{q}(v_M|p_1), \hat{q}(v_H|p_1)\} \leq 0.05$ to be close enough to the equilibrium posterior beliefs. Similarly, for the case of $p_1 = p_H$ in $M420$, any posterior beliefs $\{\hat{q}(v_L|p_1), \hat{q}(v_M|p_1), \hat{q}(v_H|p_1)\}$ such that $\hat{q}(v_L|p_1) \in [0.13, 0.17]$, $\hat{q}(v_M|p_1) \in [0.13, 0.17]$, and $\hat{q}(v_H|p_1) \in [0.66, 0.74]$ are counted as the incidences of the third-order positive selection. However, not many posterior beliefs fall within this generous definition.

**Result 3.** *Only 6.51% of the posterior beliefs (40 out of 614) are rationalized in the third-order positive selection.*

It is worth noting that the third-order positive selection leads the posterior belief to be identical to the prior belief in most cases. Thus, we must keep in mind that some of the observations may not be the result of the third-order reasoning, but the result of zero reasoning. When examining our next result, we will revisit this issue.

Hypothesis 4 regards the individual consistency. From Results 1–3, we know few subjects' posterior beliefs are consistent with the equilibrium belief. It does not, however, mean that the subjects are irrational. If their second-round price offer is consistent with the optimal price offer derived from their reported belief, then we could say that the sellers' behavior is properly driven by their incentives to maximize profit.

Overall, 66.94% (411 out of 614) of the second-round price offers were optimal from their subjective beliefs. Among those who passed the first-order positive selection, 70.22% of the second-round price offers were optimal. Among those who pass the second-order positive selection, 76.82% of the second-round price offers were optimal, indicating that higher-order reasoning on positive selection is positively associated with pricing optimality. However, among those who passed the third-order positive selection, only 55.00% of the price offers were optimal. This is because the equilibrium posterior belief $(1/3, 1/3, 1/3)$ is observationally equivalent to the most naïve subject's response—keep the prior. Once we regard only those who (1) report the posterior belief closer to the equilibrium posterior and (2) offer the optimal price given their belief in the second round as the "equilibrium sellers," there are only 22 data points (out of 614), or only 13 subjects out of 147.

**Result 4.** *A majority (66.94%) of the second-round price offers were optimal in the sense that the offer maximizes the expected profit calculated with their subjective beliefs. Higher-order reasoning on positive selection is positively associated with pricing optimality.*

Now we turn to the buyer's responses. Hypothesis 5 states that even if we observe some rejections in round 1, such rejections are more likely from the middle- or high-type buyers who receive $p_1 = p_M$ or $p_1 = p_H$. Table 5 shows the proportions of rejected price offers in round 1.

Supporting Hypothesis 5, buyers with $v_M$ or $v_H$ tend to voluntarily delay after $p_H$ more frequently than after $p_M$, and do not delay after $p_L$ at all. What is inconsistent with Hypothesis 5 is that about one third of the low-type buyers voluntarily delayed after $p_M$ and $p_H$. It implies that some low-type buyers expect that the price offer in the second round would sharply drop to $p_L$, which is both groundless and worthless. It is groundless because a rational seller who offered either $p_M$ or $p_H$ in the first round would never drop the second-round price to $p_L$. It is worthless because even in the most optimistic situation where the seller indeed offers $p_L$ in the second round, the expected payoff $\delta(v_L - p_L) = 0.8 * 60 = 48$ is strictly smaller than 50, the value of the outside option.[28]

| $v \setminus p_1$ | $p_L$ | $p_M$ | $p_H$ |
|---|---|---|---|
| $v_L$ | 9% (1/11) | 36% (76/211) | 36% (102/280) |
| $v_M$ | 0% (0/3) | 57% (137/240) | 67% (151/227) |
| $v_H$ | 0% (0/7) | 31% (72/231) | 82% (214/260) |

Table 5: Proportions of Voluntary Delay in Round 1

Each entity shows the percentage of rejections when the buyer's type corresponds to the first column, and the seller's first-round offer corresponds to the first row. The number of rejections over the number of offers is in parentheses.

**Result 5.** *Some buyers with $v_H$ or $v_M$ voluntarily delay the first-round price offer, expecting that the second-round price offer would be more favorable to them. Some low-type buyers also delay price offers where they could have been better off by exercising the outside option.*

We claim that the buyers' decisions are not driven by inequity aversion. As stated in Hypothesis 1, the low-type buyer does not have a strong reason to reject the first-round price offer, even if the buyer is averse to disadvantageous inequity.[29] Furthermore, Section 5.3 reports a supplementary experiment that incorporates the Amended Random Payment method to neutralize other-regarding motives experimentally, and the patterns observed there qualitatively replicate the main results.

Some may argue that the seller's posterior beliefs may be reflecting the buyers' (perhaps irrational) behaviors, and that may be why they did not update the posterior belief in an equilibrium way. We still claim that the discrepancies between theoretical predictions and experimental findings are mainly due to the seller's failure of positive selection reasoning. Note that even if the bargaining process moves to round 2, the seller could not observe the matched buyer's type, which means that the information presented in Table 5 is unavailable to the seller. Also, even if the sellers

[28] In the last three periods, such proportions of rejections are maintained at slightly decreased levels of 28.36% (19 out of 67) in $p_M$ and 28.92% (24 out of 83) in $p_H$, respectively. So we claim that inexperience or lack of understanding are not the primary factors behind this observation.

[29] The second-order positive selection (Hypothesis 2) is also robust to the buyer's aversion to disadvantageous inequity for the case $v_M = 90$. The reason is essentially identical to one following Hypothesis 1. Both exercising the outside option and accepting $p_M$ lead the buyer type $v_M = 90$ to obtain an advantageous payoff. Thus, any disadvantageous inequity aversion cannot counter the second-order positive selection that arises from the buyer's comparison between these two outcomes.

correctly anticipated that the proportions of voluntarily delaying in round 1 are similar to those in Table 5, the 'empirical' posterior belief, $(\hat{q}(v_L|p_1), \hat{q}(v_M|p_1), \hat{q}(v_H|p_1))$ based on such observations would be (29%, 46%, 25%) for $p_1 = p_M$ and (19%, 36%, 44%) for $p_1 = p_H$, which are depicted as blue diamonds in Figure 3. Few posterior beliefs are spotted around the empirical posterior belief, implying that it does not seem that the sellers correctly respond to the buyers' behaviors.

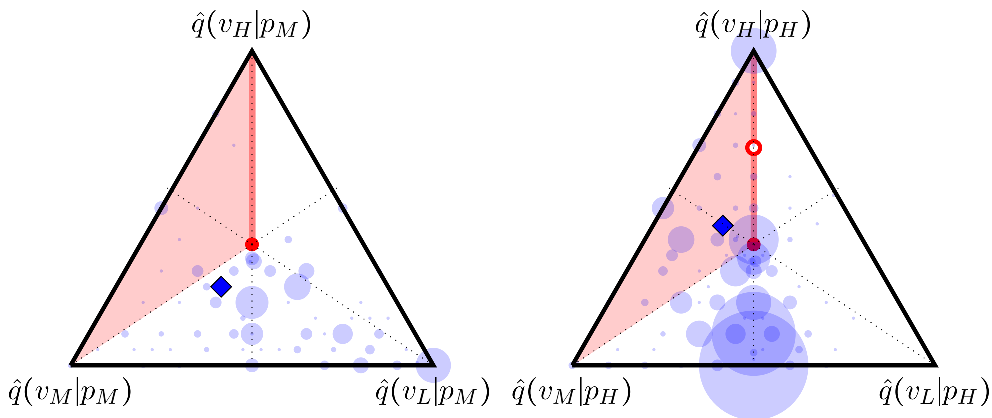


Figure 3: Observed and Empirical posterior beliefs, by $p_1$

The blue circles on the left (right) depict the reported posterior beliefs after the first-round price offers of $p_M$ ($p_H$) are met with voluntary delay. A larger circle means more observations in the center of it. The diamond shape in each simplex indicates the posterior belief consistent with the empirically observed proportions of delay in the first-round price offer.

While Figure 3 already indicates that sellers' reported beliefs do not cluster near the empirical-best-response benchmark (the diamond markers), Section 5.2 strengthens this conclusion by reporting a supplementary experiment in which the price set is restricted to $\{p_M, p_H\}$, mathematically eliminating any coherent rationale for assigning positive probability to the low type after a first-round rejection. Even in that environment, sellers continue to hold a positive posterior belief on the low type, confirming that the failure of positive selection reflects a genuine inability to apply the iterative reasoning rather than a sophisticated response to anticipated buyer behavior.

In addition, we count how many first-round price offers are equal to the equilibrium offer. Overall, 64.29% of the offers (945 out of 1,470) in the first round were the profit-maximizing price. Figure 4 shows the proportion of each price per treatment, and black bars indicate the proportion of the optimal price.

As barely noticeable, few sellers offer $p_1 = 10$ in all treatments. While the optimal price ($p_H$ and $p_M$, respectively) was most frequently offered in $M90$ and $M420$, $p_M$ instead of the optimal price $p_H$ was most frequently offered in $M240$. Does it mean that the sellers in $M240$ were less rational? We doubt this. It is more plausible that the experiment participants are inclined to offer the middle price if that middle one is not too far from what they expect to be optimal.

Lastly, we explore whether subjects' failure of positive selection reasoning attenuates with expe-

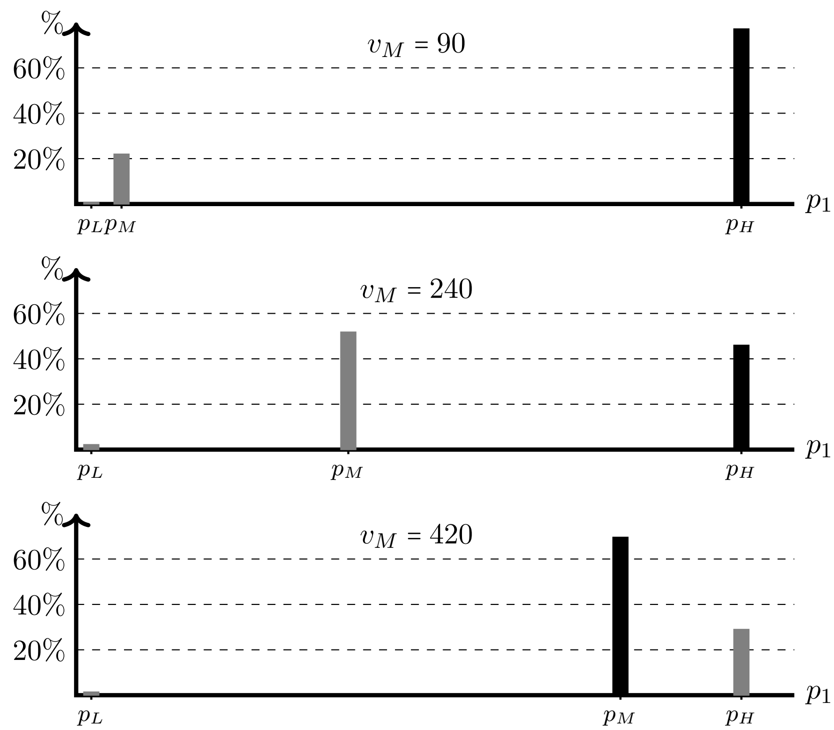


Figure 4: Optimal pricing or targeting the middle

Each bar depicts the percentage of the first-round offers. The black bar indicates the optimal price offer given $v_M$.

rience. The average first-round price offer remains stable across all ten periods, with no statistically significant time trend ($p = 0.520$). Examining sellers' reported posterior beliefs, the misspecified component proxied by $\hat{q}(v_L|p_1)$—which theory predicts to be less than the uniform prior—declines significantly over the first five periods ($p = 0.021$), suggesting some directional learning at the outset. However, this downward trend halts thereafter, with $\hat{q}(v_L|p_1)$ persisting at approximately 32% across the last five periods ($p = 0.472$), still far above the equilibrium benchmark. Figure 5 juxtaposes the reported posterior beliefs from the first and last five periods, and the bulk of the distribution remains in the bottom region of the simplex in both panels, with only a modest movement toward the equilibrium prediction. The individual consistency between sellers' subjective beliefs and their second-round price offers documented in Result 4 also exhibits no significant change over time ($p = 0.992$). Taken together, these patterns suggest that while subjects make some incremental belief adjustment early on, the cognitive hurdle posed by higher-order positive selection reasoning is not readily surmounted through short-run repetition.

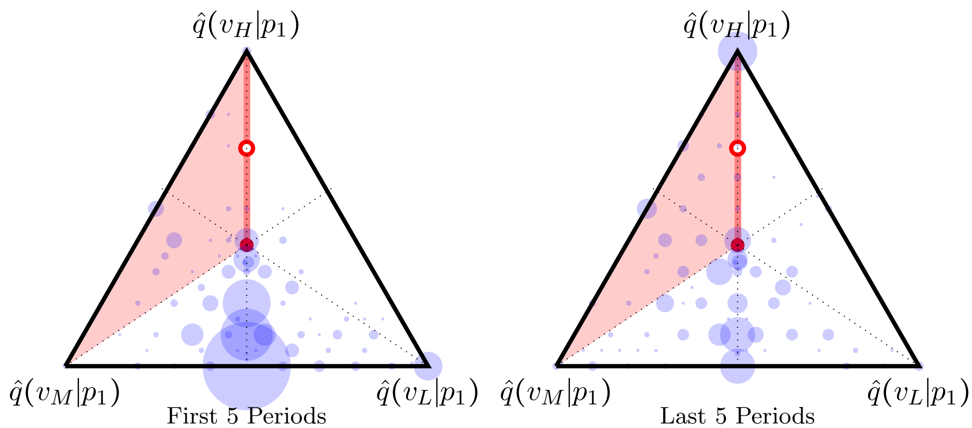


Figure 5: Observed posterior beliefs, by first and last 5 periods

The blue circles on the left (right) depict the reported posterior beliefs after the first-round price offers are rejected in the first (last) 5 periods. A larger circle means more observations in the center of it.

# 5 Supplementary Experiments

The main results invite four natural alternative explanations for why subjects fail at positive selection. We address each through a dedicated supplementary experiment. Sections 5.1 and 5.2 report two supplementary experiments tied directly to Results 1 and 5, respectively. Sections 5.3 and 5.4 summarize two further supplementary experiments designed to address concerns about other-regarding preferences and the seller's general pricing competence. Full experimental procedures and additional analyses for all four are reported in Appendix C and the Online Appendix.

### 5.1 Is the failure due to a general inability at strategic reasoning?

A natural concern about Result 1 is that it might reflect not a specific failure of positive selection but a more general inability of subjects to engage in any kind of equilibrium reasoning about how the demand pool composition changes after a rejection. To address this, we conduct a supplementary experiment that removes the buyer's outside option entirely. Without an outside option, the equilibrium reasoning requires negative selection: a delay in the negotiation indicates that the remaining demand pool consists of relatively lower types, since some high-type buyers are likely to accept the first-round offer. Negative selection thus serves as a mirror image of positive selection, which has the same belief-updating discipline but in the opposite direction.

The theoretical predictions for the no-outside-option environment are derived in Appendix B. We use the parameter values $v_L = 20$, $v_M = 190$, $v_H = 450$, $\Delta = 10$, and $\delta = 0.8$, calibrated so that the gain from trading is identical to that in our main treatment M240. In equilibrium, the seller offers $p_1 = p_H = 440$ in the first round, and the high-type buyer randomizes by accepting this offer with probability 0.308. If the offer is rejected, the seller's on-equilibrium-path posterior is $(\hat{q}(v_L), \hat{q}(v_M), \hat{q}(v_H)) = (0.371, 0.371, 0.257)$, a clear instance of negative selection, since the posterior places more weight on the low and middle types and less weight on the high type relative to the prior.

Our experimental findings are broadly consistent with these predictions. The first-round offer $p_H$ was selected in 48.60% of all observations, and the average offer (305.19) is comparable to the average offer in M240 (295.86). Buyer responses are also consistent with the theory: most middle- and low-type buyers rejected the first-round offer. After rejection, 97.8% of the sellers either decreased or maintained the second-round price, indicating that "obvious" mistakes, that is, raising the price after rejection, are rarely observed. Most importantly, Figure 6 displays the sellers' reported posterior beliefs after a first-round rejection. The majority of beliefs cluster in the region consistent with negative selection, and many sellers update even more aggressively than the theory predicts (the red hollow circle in the figure indicating the on-equilibrium-path posterior).

An eyeball comparison between Figure 6 and the main experiment's Figure 2 reveals that the overall distribution of reported posterior beliefs looks broadly similar across the two environments, with the bulk of observations concentrated near the bottom of the simplex in both cases. The crucial difference, however, is interpretive: in the no-outside-option environment of Figure 6, this concentration is consistent with negative selection and thus with equilibrium reasoning, whereas in the main experiment of Figure 2 the same region of the simplex represents a failure of positive selection.

These observations indicate that subjects are capable of the basic strategic reasoning underlying belief updating after a rejection. The failure of positive selection in our main experiment, therefore, is not attributable to a general inability to engage in such reasoning. It appears, rather, to be a

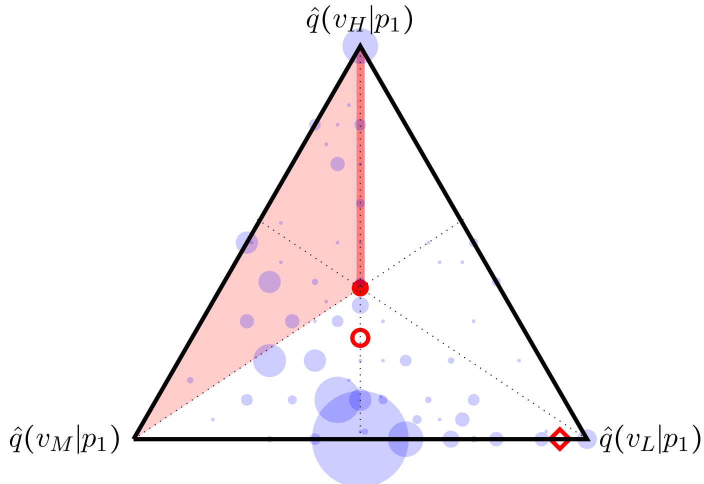


Figure 6: Observed posterior beliefs, without outside option

The blue circles depict the reported posterior beliefs after the first-round price offers are rejected. A larger circle means more observations in the center of it. A hollow red circle indicates the equilibrium posterior belief, and a hollow red diamond indicates the posterio belief when $p_1 = p_M$.

specific inability to execute the harder, higher-order reasoning that positive selection requires, where the seller must combine reasoning about each buyer type's incentive to take the outside option (rather than to accept or reject) with the resulting upward shift in the demand pool's composition.

## 5.2 Is the failure due to sellers' concern about buyer irrationality?

Another concern about Result 5 is that sellers may place significant weight on the low type remaining in the second round not because they fail at positive selection per se, but because they correctly anticipate the irrational behavior of some buyers, specifically, the low-type buyers who reject $p_M$ and $p_H$ instead of taking the outside option (Table 5). Under this alternative interpretation, sellers' posterior beliefs are a rational response to observed empirical behavior rather than a failure of higher-order reasoning. We have already argued in Section 4 that this interpretation is unlikely on the basis of Figure 3: if sellers were genuinely best-responding to the empirical rejection rates, their reported beliefs should cluster near the blue diamond in each simplex, which they do not. The supplementary experiment described in this subsection addresses the concern more directly by constructing an environment which reduces the scope for this alternative interpretation sharply.

To rule out this alternative, we conduct a second supplementary experiment that the first step of iterated dominance is more transparent. We restrict the seller's price set to $\{p_M, p_H\}$, removing $p_L$ altogether. Under this restriction, a low-type buyer cannot afford any feasible second-round price: taking the outside option is the *unique* choice for the low type, not merely a strategically dominant one. As a consequence, even a seller who entertains substantial uncertainty about the buyer's (ir)rationality has no coherent basis for assigning positive probability to the low type remaining in the demand pool after a first-round rejection.

The data confirm that subjects largely understand this restriction. 88% of the low-type buyers (125 out of 142) took the outside option in the first round, while only 2% of the high-type buyers (3 out of 154) did so, indicating that low types correctly recognize that delaying is futile and high types correctly recognize that the outside option is unattractive. Despite this, many of sellers' reported posterior beliefs violate the first-order positive selection. Figure 7 displays the 74 posterior beliefs reported by sellers whose first-round offers were rejected. The reported posterior beliefs cluster predominantly along the bottom of the unit simplex, with a substantial share assigning the low type a probability above its prior probability of 1/3, contrary to the first-order positive-selection $\hat{q}(v_L|p_1) \leq 1/3$.

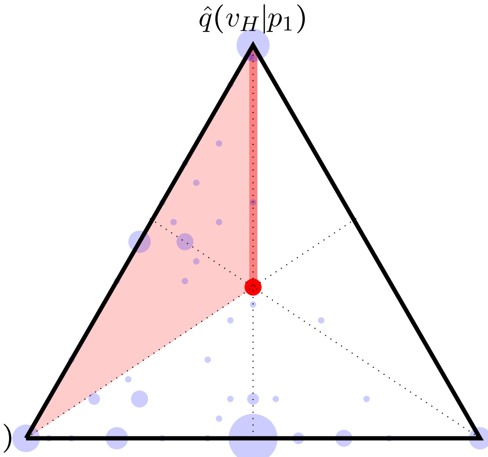


Figure 7: Observed posterior beliefs, without $p_L$ option

This pattern provides direct evidence against sellers' anticipation of buyer irrationality as the primary explanation for the failure of positive selection. Even when the first step of iterated dominance is made especially transparent, substantially weakening the channel through which a "rational concern about buyer irrationality" could matter, reported beliefs still frequently fail to downweight the low type as first-order positive selection requires. The persistence of this belief-updating failure suggests that sellers have difficulty carrying out the iterative reasoning at the heart of positive selection, rather than merely adjusting their beliefs to anticipated buyer behavior.

### 5.3 Is the failure due to other-regarding preferences?

A third concern is that the failure of positive selection might be an artifact of other-regarding preferences rather than a cognitive failure. On the seller's side, inequity aversion could deter sellers from offering the high equilibrium price even when they understand its profitability. On the buyer's side, advantageous-inequity aversion could lead some buyers to voluntarily delay, despite recognizing that doing so is suboptimal. We argued in Sections 3 and 4 that the design of our main experiment, with three discrete prices spaced far apart and an always-available outside option, leaves limited room for such preferences to operate. We further validate this argument experimentally.

We conduct three additional sessions, one per treatment, with a modified design that incorporates the Amended Random Payment (ARP) method (Lim and Xiong, 2021). Under ARP, each participant's final payment is determined by an independently drawn random match, and participants are told whether the current match determines their paired participant's payment. Crucially, in the 90% of matches that do not determine the paired participant's payment, decisions cannot affect the other participant's earnings, which is known to neutralize other-regarding motives effectively (Nishimura and Hanaki, 2024).

If other-regarding preferences were driving our main results, we would expect rejection rates to fall and posterior beliefs to shift toward equilibrium under ARP. Neither pattern emerges. Rejection rates remain substantially above the theoretical prediction of zero. For instance, 73% of middle-type buyers reject $p_M$ in M420 under ARP, compared to a comparable rate in the main experiment. Figure 8 displays the sellers' reported posterior beliefs under ARP. As in the main experiment, the vast majority of observations lie outside the region consistent with first-order positive selection.

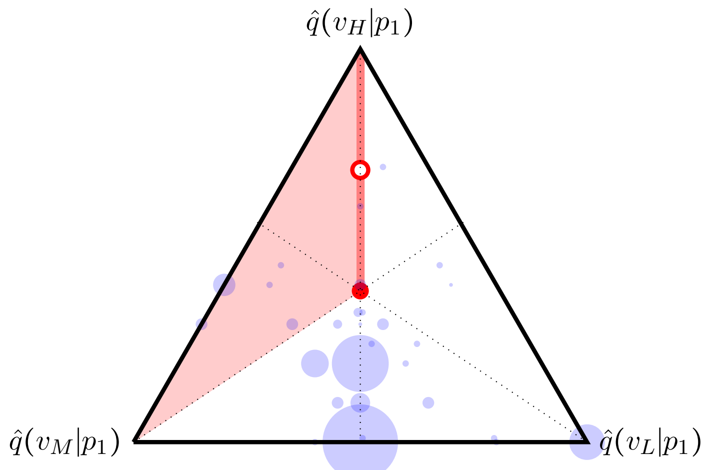


Figure 8: Observed posterior beliefs: ARP

These observations confirm that other-regarding preferences are not the primary driver of the failure of positive selection. The same qualitative patterns—substantial rejection rates and posterior beliefs inconsistent with first-order positive selection—emerge even in an environment in which other-regarding motives have been experimentally neutralized.

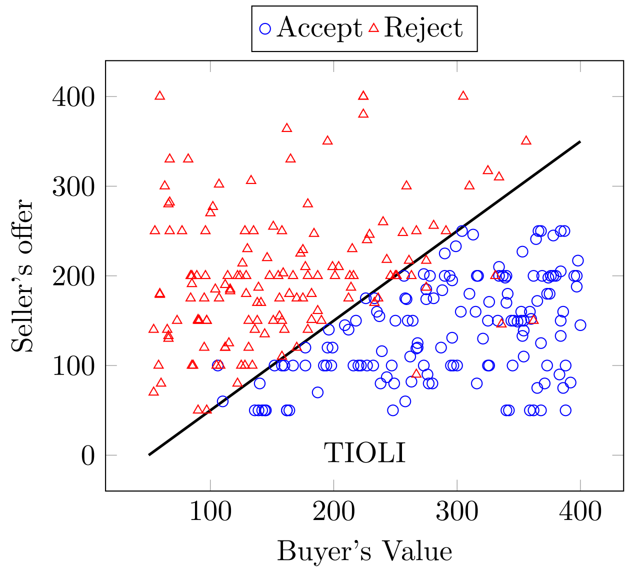


Figure 9: Scatterplot of Buyer's Value and Seller's Offer in the TIOLI experiment

### 5.4 Is the failure due to sellers' limited understanding of optimal pricing?

A fourth concern is that the failure of positive selection might be a symptom of a more basic limitation: sellers may simply not know how to compute a profit-maximizing price, even in the absence of strategic complications. If so, the failure we document would have little to do with positive selection per se and would instead reflect a generic difficulty with monopoly pricing under uncertainty. Relatedly, sellers might price below the static optimum because they anticipate that buyers would reject "unfair" offers, entangling pricing competence with fairness considerations.

To address both concerns, we conduct four additional sessions of an ultimatum bargaining experiment with private information and an outside option. In each match, a privately-informed buyer with value drawn uniformly from $[50, 400]$ faces a single take-it-or-leave-it (TIOLI) offer from a randomly paired seller, and the buyer's outside option is worth 50. The TIOLI design strips away the dynamic-screening structure of our main experiment, isolating the seller's pricing decision from any need to reason about continuation play, posterior beliefs, or positive selection.

Under our parameter values, the offer that maximizes the seller's expected profit is 175. The average TIOLI offer is 166.92, which is not statistically distinguishable from 175 ($p = 0.391$, two-sided $t$-test). Figure 9 plots the buyer's value against the seller's offer, with the buyer's indifference line between accepting and exercising the outside option. Acceptance and rejection are well separated by this indifference line: buyers accept when accepting yields more than the outside option, and they reject otherwise, even when accepting yields a substantial payoff inequity in favor of the seller.

Two implications follow. First, sellers on average understand the profit-maximizing offer in a static pricing problem. Their failure to charge the equilibrium price in the main experiment thus cannot be attributed to a generic difficulty with monopoly pricing. Second, buyers accept inequitable offers when accepting is in their material self-interest, suggesting that sellers in the main experiment have no reason to anticipate widespread fairness-motivated rejections.

### 5.5 Summary of Supplementary Experiments

Taken together, the four supplementary experiments rule out the most natural alternative explanations for our main findings. The failure of positive selection is not a manifestation of a general inability at strategic reasoning (5.1), of sellers' anticipation of irrational buyer behavior (5.2), of other-regarding preferences (5.3), or of sellers' limited understanding of optimal pricing (5.4). Each candidate explanation, by itself, would predict patterns in the data—fewer rejections, posterior beliefs that rationally anticipate buyers' irrational rejections, weaker effects under ARP, or systematically suboptimal TIOLI offers—that do not appear. What remains is the interpretation we propose: subjects find positive selection genuinely difficult to execute, even when the underlying belief-updating discipline is within their reach (5.1) and even when the first step of the iterated argument is mathematically forced (5.2). The failure documented in our main experiment is, in this sense, a clean diagnosis of a specific cognitive limitation rather than a confound.

## 6 Conclusion

In this paper, we experimentally examine how the idea of positive selection unfolds in a two-round bargaining game with one-sided, incomplete information. Inconsistent with theory, a substantial proportion of first-round price offers are rejected when an outside option is available. The reported beliefs from the sellers after a price offer is rejected confirm that the higher-order positive selection rarely takes place in the lab: Even in our generous definition, only about 7% of the posterior beliefs are considered as the result of the third-order positive selection. Worse yet, about half of those with the "correct" posterior belief were the most naïve ones who merely stuck to the prior and could not find the profit-maximizing price. Overall, this failure of the inductive process of positive selection leads the monopoly seller to earn profits much less than what the theory predicts.

The main contribution of this paper is to decipher the level at which positive selection reasoning fails. This is beyond merely examining the validity of the theoretical predictions of the game, and we believe it is more important for counterfactual analyses and policy suggestions. We found a clear rejection of the theoretical predictions, and our experimental design allows us to enunciate the drivers of the discrepancies between experimental data and theoretical predictions. We hope our method of experimentally stripping down the several layers of rational reasoning to be used in various contexts whose equilibrium predictions are derived from multiple thought processes.

Our central finding is that the higher-order iterated reasoning required for positive selection is cognitively demanding and prone to systematic failure. The accompanying belief and rejection patterns are consistent with subjects defaulting to a simpler conjecture, namely that delay is rewarded with a lower price. This is the correct logic under negative selection but precisely the wrong logic once an outside option induces positive selection. This interpretation also clarifies why the two selection regimes carry opposite policy implications. Under negative selection, the Coasean force already works in buyers' favor, so restricting access to outside options is what protects consumer surplus. Under positive selection, the prescription inverts: the same outside option is what converts surplus into seller profit. Our results suggest this inversion, while theoretically sound, rarely materializes in practice, because the reasoning that sustains it is rarely executed. A market designer who assumes positive selection will operate as theorized may therefore substantially overstate the profit a monopolist can extract. We also note that our setting, a bilateral monopoly with a fixed, exogenous outside option, abstracts from broader market complexities: in many real-world environments, market structure involves competition, and outside options may respond endogenously to the monopolist's actions. How these boundedly rational tendencies interact with competitive and endogenous outside options is an important boundary condition for our policy claims.

More broadly, the failure we document speaks to unraveling, a fundamental mechanism in many settings beyond bargaining, including information disclosure. Our findings suggest that the observed failures of unraveling in both laboratory and real-world settings (Jin et al., 2021; Brown et al., 2012) may stem primarily from a lack of strategic sophistication among participants. Because our analysis locates the breakdown at a specific step of the iterated reasoning, it raises the question of whether an intervention that completes the initial step on participants' behalf could trigger the remainder of the process. To our knowledge, no prior work has examined the responsiveness of unraveling to such interventions at the initial stage, and we consider this an intriguing direction for future research.

## References

**Abeler, Johannes, Daniele Nosenzo, and Collin Raymond**, "Preferences for truth-telling," *Econometrica*, 2019, *87* (4), 1115–1153.

**Bereby-Meyer, Yoella and Muriel Niederle**, "Fairness in bargaining," *Journal of Economic Behavior & Organization*, 2005, *56* (2), 173–186.

**Board, Simon and Marek Pycia**, "Outside options and the failure of the Coase conjecture," *American Economic Review*, 2014, *104* (2), 656–71.

**Brandts, Jordi and Gary Charness**, "The strategy versus the direct-response method: a first survey of experimental comparisons," *Experimental Economics*, 2011, *14*, 375–398.

**Brown, Alexander L., Colin F. Camerer, and Dan Lovallo**, "To Review or Not to Review? Limited Strategic Thinking at the Movie Box Office," *American Economic Journal: Microeconomics*, May 2012, *4* (2), 1–26.

**Burchardi, Konrad B. and Stefan P. Penczynski**, "Out of your mind: Eliciting individual reasoning in one shot games," *Games and Economic Behavior*, 2014, *84*, 39–57.

**Burdea, Valeria and Jonathan Woon**, "Online belief elicitation methods," *Journal of Economic Psychology*, 2022, *90*, 102496.

**Catonini, Emiliano**, "The failure of Coase conjecture with outside options: a rationalizability approach," 2022. Working paper.

**Chang, Dongkyu and Jong Jae Lee**, "Price Skimming: Commitment and Delay in Bargaining with Outside Option," *Journal of Economic Theory*, 2022, *205*, 105528.

**_ , Duk Gyoo Kim, and Wooyoung Lim**, "Positive and negative selection in bargaining," *Journal of Economic Theory*, 2026, *235*, 106190.

**Chen, Daniel L., Martin Schonger, and Chris Wickens**, "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88–97.

**Chen, Jing and Silvio Micali**, "The Order Independence of Iterated Dominance in Extensive Games," *Theoretical Economics*, 2013, *8* (1), 125–163.

**Coase, Ronald H.**, "Durability and Monopoly," *The Journal of Law and Economics*, 1972, *15* (1), 143–149.

**Danz, David, Lise Vesterlund, and Alistair J. Wilson**, "Belief Elicitation and Behavioral Incentive Compatibility," *American Economic Review*, September 2022, *112* (9), 2851–2883.

**Demichelis, Stefano and Jörgen W Weibull**, "Language, meaning, and games: A model of communication, coordination, and evolution," *American Economic Review*, 2008, *98* (4), 1292–1311.

**Fanning, Jack**, "Outside options, reputations, and the partial success of the Coase conjecture," *Econometrica*, 2026, *94* (3), 877–910.

**Fehr, Ernst and Klaus M. Schmidt**, "A theory of fairness, competition, and cooperation," *The Quarterly Journal of Economics*, 1999, *114* (3), 817–868.

**Fudenberg, Drew and Jean Tirole**, *Game Theory*, The MIT Press, 1991.

—, **David Levine, and Jean Tirole**, "Infinite Horizon Models of Bargaining with One-sided Incomplete Information," in A. E. Roth, ed., *Bargaining with Incomplete Information*, London/New York: Cambridge University Press, 1985, pp. 73–98.

**Gneezy, Uri, Agne Kajackaite, and Joel Sobel**, "Lying aversion and the size of the lie," *American Economic Review*, 2018, *108* (2), 419–453.

**Gul, Faruk, Hugo Sonnenschein, and Robert Wilson**, "Foundations of Dynamic Monopoly and the Coase Conjecture," *Journal of Economic Theory*, 1986, *39* (1), 155–190.

**Ho, Teck-Hua and Xuanming Su**, "A Dynamic Level-k Model in Sequential Games," *Management Science*, 2013, *59* (2), 452–469.

**Hwang, Ilwoo and Fei Li**, "Transparency of outside options in bargaining," *Journal of Economic Theory*, 2017, *167*, 116–147.

**Jin, Ginger Zhe, Michael Luca, and Daniel Martin**, "Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure," *American Economic Journal: Microeconomics*, May 2021, *13* (2), 141–173.

**Kim, Duk Gyoo**, "“One Bite at the apple”: Legislative bargaining without replacement," *Journal of Economic Psychology*, 2023, *95*, 102589.

**Kneeland, Terri**, "Identifying Higher-Order Rationality," *Econometrica*, 2015, *83* (5), 2065–2079.

**Lim, Wooyoung and Siyang Xiong**, "Does jump bidding increase sellers' revenue? Theory and experiment," *Journal of Economic Behavior & Organization*, 2021, *189*, 84–110.

**Nishimura, Takeshi and Nobuyuki Hanaki**, "On the performance of the Crémer-McLean auction: An experiment," Technical Report, ISER Discussion Paper 2024.

**Pearce, David G.**, "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 1984, *52* (4), 1029–1050.

**Roth, Alvin E. and J. Keith Murnighan**, "Equilibrium behavior and repeated play of the prisoner's dilemma," *Journal of Mathematical psychology*, 1978, *17* (2), 189–198.

**Shimoji, Makoto and Joel Watson**, "Conditional Dominance, Rationalizability, and Game Forms," *Journal of Economic Theory*, 1998, *83* (2), 161–195.

**Tirole, Jean**, "From Bottom of the Barrel to Cream of the Crop : Sequential Screening with Positive Selection," *Econometrica*, 2016, *84* (4), 1291–1343.

# Online Appendices

## A Details Omitted in Section 2

This appendix includes the results and proofs omitted in Section 2. Throughout the appendix, let $V = \{v_1, \ldots, v_N\}$, with $v_1 < \cdots < v_N$, and suppose $q(v_k) > 0$ for every $k = 1, \ldots, N$. Recall that $u(v_k) = v_k - w$ and $\bar{p}(v_k) = \max\{p \in \mathcal{P} : u(v_k) > p\}$. As in the main text, let $\Delta := \max_{v \in V}\{u(v) - \bar{p}(v)\}$ and define $K^* := \max\{k \in \{1, \ldots, N\} : \bar{p}(v_k) \leq p^*\}$, so that $v^* = v_{K^*}$.

### A.1 $p_1$-Continuation Game and Conditional Dominance

We first formally define the $p_1$-continuation game. Shimoji and Watson (1998) focus on extensive-form games without moves of Nature, whereas our model includes both Nature's initial draw of the buyer's type and its intervention at the end of round 1. We adapt their augmented-normal-form framework to our $p_1$-continuation games by (i) representing each buyer type as a separate buyer-agent, as in the type-agent representation of a Bayesian game, (ii) incorporating the initial type draw and Nature's intervention into the expected payoff functions, and (iii) introducing normal-form restrictions corresponding to the relevant information sets of the $p_1$-continuation game.

Also, note that, after any second-round price $p_2 \in \mathcal{P}$, buyer-agent $k$ accepts if $u(v_k) > p_2$ and exercises the outside option otherwise. Indeed, rejecting both alternatives is strictly inferior because $w > 0$, and Assumption (A1)-($i$) rules out indifference between accepting and exercising the outside option. Thus, every alternative second-round response is strictly dominated in the corresponding information-set restriction. We may therefore eliminate these responses before applying the deletion procedures. After their elimination, we incorporate the payoff from the unique second-round response into the payoff functions and focus on the buyer-agent's response to $p_1$.

*Normal Form.* We represent the game as a finite normal-form game among $N + 1$ players indexed by $I = \{0, 1, \ldots, N\}$. Player 0 is the seller. For each $k \geq 1$, player $k$ is the buyer-agent corresponding to type $v_k$. Fix $p_1 \in \mathcal{P}$. The $p_1$-continuation game begins immediately after $p_1$ has been offered and before the buyer responds. The buyer's type remains distributed according to the prior $q$. Let $\mathcal{B}_1 := \{A, O, D\}$ denote the buyer's possible responses to $p_1$. In view of the preceding observation, we omit the buyer-agent's second-round response from her strategy. For either $\epsilon > 0$ or $\epsilon = 0$, a pure strategy of player 0 is a second-round price $p_2 \in S_0^\epsilon := \mathcal{P}$, and a pure strategy of buyer-agent $k$ is a response $a_k \in S_k^\epsilon := \mathcal{B}_1$ to the fixed first-round price $p_1$. Let $S^\epsilon := \prod_{i \in I} S_i^\epsilon$ denote the resulting finite strategy space. Although this strategy space is the same for $\epsilon > 0$ and $\epsilon = 0$, we retain the superscript because the payoff functions and information-set restrictions depend on $\epsilon$.

Once a pure-strategy profile $s \in S^\epsilon$ is fixed, all expected payoffs are well defined. Player 0's payoff is the seller's expected payoff, integrating over the prior $q$, Nature's intervention probability

$\epsilon$, and the discount factor $\delta$. Player $k$'s payoff is type $v_k$'s expected payoff conditional on the buyer's type being $v_k$. Let $U^{p_1,\epsilon} = (U_0^{p_1,\epsilon}, \ldots, U_N^{p_1,\epsilon})$ denote these payoff functions, extended to mixed strategies in the usual way. We write $G^\epsilon(p_1) \coloneqq (I, S^\epsilon, U^{p_1,\epsilon})$ for the normal form of the $p_1$-continuation game.

*Information-Set Restrictions.* We augment $G^\epsilon(p_1)$ with the information-set restrictions of the $p_1$-continuation game. We again define the restriction collections separately for $\epsilon > 0$ and $\epsilon = 0$. In either case, the restriction collections do not depend on the fixed first-round price $p_1 \in \mathcal{P}$.

First, suppose $\epsilon > 0$. Under every buyer-strategy profile, round 2 is reached with positive probability: it is reached for sure if the realized buyer type chooses $D$, and Nature may force continuation if the buyer chooses $A$ or $O$. Moreover, the seller observes neither the buyer's type nor whether continuation was voluntary or forced by Nature. Hence, the seller's unique second-round information-set restriction is the entire strategy space $S^\epsilon$. Each buyer-agent's first-round information-set restriction is also $S^\epsilon$, because she chooses her response to the fixed price $p_1$ before observing the seller's second-round price $p_2$. Therefore, each player $i$'s collection of restrictions is $\mathcal{A}_i^\epsilon = \{S^\epsilon\}$. We write $\mathcal{A}^\epsilon \coloneqq (\mathcal{A}_0^\epsilon, \mathcal{A}_1^\epsilon, \ldots, \mathcal{A}_N^\epsilon)$.

Now suppose that $\epsilon = 0$. The seller reaches round 2 only if the realized buyer type voluntarily delays. Let

$$D_{-0}^0 \coloneqq \{s_{-0} \in S_{-0}^0 : s_k = D \text{ for some } k\}.$$

Thus, $D_{-0}^0$ is the set of buyer-agent strategy profiles consistent with some buyer type voluntarily delaying. The seller's restriction collection is $\mathcal{A}_0^0 = \{S_0^0 \times D_{-0}^0\}$, where $S_0^0 \times D_{-0}^0$ is the restriction corresponding to the seller's second-round information set. Upon reaching this information set, the seller knows that the buyer voluntarily delayed but does not observe her type. Similar to the case $\epsilon > 0$, each buyer-agent's first-round information-set restriction is the entire strategy space $S^0$, so $\mathcal{A}_k^0 = \{S^0\}$ for each buyer-agent $k \geq 1$. We write $\mathcal{A}^0 \coloneqq (\mathcal{A}_0^0, \mathcal{A}_1^0, \ldots, \mathcal{A}_N^0)$.

Note that, when $\epsilon > 0$, each player has the single restriction $S^\epsilon$, so that $X \cap Y = Y$ for every product set $Y \subseteq S^\epsilon$; conditional dominance then coincides with ordinary strict dominance in the reduced $p_1$-continuation game. We retain the conditional-dominance terminology to treat the perturbed and unperturbed cases within a common framework. The restriction is substantive when $\epsilon = 0$, where $S_0^0 \times D_{-0}^0$ is a proper subset of the strategy space.

*Conditional-Dominance Deletion.* Following Shimoji and Watson (1998), we define conditional dominance as follows. Let $Y = \prod_{i \in I} Y_i \subseteq S^\epsilon$ be a nonempty product set of remaining strategies, and let $X = X_i \times X_{-i} \in \mathcal{A}_i^\epsilon$ be an information-set restriction for player $i$. A strategy $s_i \in X_i \cap Y_i$ is *strictly dominated in* $X \cap Y$ if $X_{-i} \cap Y_{-i} \neq \varnothing$ and there exists $\alpha_i \in \Delta(X_i \cap Y_i)$ such that

$$U_i^{p_1,\epsilon}(\alpha_i, s_{-i}) > U_i^{p_1,\epsilon}(s_i, s_{-i}) \qquad \forall s_{-i} \in X_{-i} \cap Y_{-i}.$$

A strategy $s_i$ is *conditionally dominated on* $(\mathcal{A}_i^\epsilon, Y)$ if there exists $X \in \mathcal{A}_i^\epsilon$ such that $s_i$ is strictly dominated in $X \cap Y$.

### A.2 Proof of Proposition 1

Fix $p_1 \in \mathcal{P}$ throughout this subsection. The first lemma formalizes the recursive step of the low-to-high deletion.

**Lemma A.1** (Recursive low-to-high deletion)**.** *Suppose Assumptions* (A1) *and* (A2) *hold. If* $\epsilon > 0$*, suppose in addition that Assumption* (A3) *holds. Fix* $k \in \{1, \dots, N\}$ *if* $\epsilon = 0$*, and* $k \in \{1, \dots, K^*\}$ *if* $\epsilon > 0$*. Let* $Y = \prod_{i \in I} Y_i \subseteq S^\epsilon$ *be a nonempty product set such that* $D \notin Y_i$ *for every* $i < k$*,* $D, O \in Y_k$*, and* $\bar{p}(v_k) \in Y_0$*. Then:*

*(i) Every* $p \in Y_0$ *such that* $p < \bar{p}(v_k)$ *is conditionally dominated on* $(\mathcal{A}_0^\epsilon, Y)$*.*

*(ii) Let* $\widetilde{Y}_0 := \{p_2 \in Y_0 : p_2 \geq \bar{p}(v_k)\}$ *and* $\widetilde{Y} := \widetilde{Y}_0 \times \prod_{i=1}^N Y_i$*. Then* $D$ *is conditionally dominated for buyer-agent* $k$ *on* $(\mathcal{A}_k^\epsilon, \widetilde{Y})$*.*

Here, $Y$ denotes the set of strategies remaining immediately before the $k$-th step of the low-to-high deletion chain. The condition $D \notin Y_i$ for every $i < k$ means that voluntary delay has already been eliminated for types $v_1, \dots, v_{k-1}$. The conditions $D, O \in Y_k$ mean that type $v_k$'s voluntary delay remains to be eliminated and that the outside option remains available as the dominating action. Finally, $\bar{p}(v_k) \in Y_0$ ensures that the seller can use $\bar{p}(v_k)$ to dominate lower second-round prices.

*Proof.* For part (i), fix any $p \in Y_0$ such that $p < \bar{p}(v_k)$. Consider any opponent strategy profile consistent with the seller's second-round information-set restriction and the remaining strategy set $Y$. Thus, take $s_{-0} \in D_{-0}^0 \cap Y_{-0}$ if $\epsilon = 0$, and $s_{-0} \in Y_{-0}$ if $\epsilon > 0$. Let

$$J(s_{-0}) := \{j \in \{k, \dots, N\} : s_j = D\}.$$

Since voluntary delay has already been eliminated for types below $v_k$, only types in $J(s_{-0})$ may voluntarily delay. Moreover, if $\epsilon = 0$, the set $J(s_{-0})$ is nonempty because the seller's second-round information set is reached only after voluntary delay.

The seller's payoff gain from replacing $p$ with $\bar{p}(v_k)$ is

$$U_0^{p_1,\epsilon}(\bar{p}(v_k), s_{-0}) - U_0^{p_1,\epsilon}(p, s_{-0}) = \delta\epsilon\big[\Pi(\bar{p}(v_k)) - \Pi(p)\big] + \delta(1-\epsilon)[\bar{p}(v_k) - p] \sum_{j \in J(s_{-0})} q(v_j).$$

Every type in $J(s_{-0})$ accepts both prices because $u(v_j) \geq u(v_k) > \bar{p}(v_k) > p$. Hence the second term is weakly positive. If $\epsilon = 0$, the first term on the right hand side of the equation is zero, while the second term is strictly positive because $J(s_{-0})$ is nonempty. If $\epsilon > 0$, then $p < \bar{p}(v_k) \leq p^*$, so Assumption (A3) implies $\Pi(\bar{p}(v_k)) > \Pi(p)$ and the first term is strictly positive. Thus, in either case, replacing $p$ with $\bar{p}(v_k)$ strictly increases the seller's payoff. Because $\bar{p}(v_k) \in Y_0$, the price $p$ is conditionally dominated on $(\mathcal{A}_0^\epsilon, Y)$.

For part (ii), take any $p_2 \in \widetilde{Y}_0$, and let $c(v_k, p_2) := \delta \max\{w, v_k - p_2\}$ denote buyer-agent $k$'s continuation payoff. Since $p_2 \geq \bar{p}(v_k)$,

$$c(v_k, p_2) \leq \delta \max\{w, v_k - \bar{p}(v_k)\} = \delta\big[w + u(v_k) - \bar{p}(v_k)\big] \leq \delta(w + \Delta) < w,$$

where the equality follows from $\bar{p}(v_k) < u(v_k) = v_k - w$, and the last inequality follows from Assumption (A2). Choosing $D$ yields $c(v_k, p_2)$. By contrast, choosing $O$ yields $(1-\epsilon)w + \epsilon c(v_k, p_2)$: the buyer obtains $w$ if Nature does not intervene and receives the same continuation payoff $c(v_k, p_2)$ if Nature forces continuation. Hence,

$$U_k^{p_1,\epsilon}(p_2, O) - U_k^{p_1,\epsilon}(p_2, D) = (1-\epsilon)\big[w - c(v_k, p_2)\big] > 0 \qquad \forall p_2 \in \widetilde{Y}_0,$$

where the strict inequality uses $\epsilon < 1$. Because $D, O \in Y_k$, the outside option is available as a dominating strategy. Thus, $D$ is strictly dominated in $S^\epsilon \cap \widetilde{Y} = \widetilde{Y}$ and hence conditionally dominated on $(\mathcal{A}_k^\epsilon, \widetilde{Y})$. $\square$

Now we are ready to prove Proposition 1. Fix $p_1 \in \mathcal{P}$. Let $K^\dagger := N$ if $\epsilon = 0$ and $K^\dagger := K^*$ if $\epsilon > 0$. We apply Lemma A.1 successively for $k = 1, \ldots, K^\dagger$. Start from the full strategy set $Y^0 := S^\epsilon$. At the first step, the conditions of Lemma A.1 hold for $k = 1$: the condition on lower types is vacuous, $D, O \in Y_1^0$, and $\bar{p}(v_1) \in Y_0^0 = \mathcal{P}$. Part (i) removes every second-round price below $\bar{p}(v_1)$, and part (ii) then eliminates type $v_1$'s voluntary delay. This yields the first-order deletion of voluntary delay.

Now fix $k \in \{2, \ldots, K^\dagger\}$ and suppose the first $k-1$ steps of the low-to-high deletion chain have been completed. Let $Y^{k-1} = \prod_{i \in I} Y_i^{k-1}$ denote the remaining strategy set. By construction, $D \notin Y_i^{k-1}$ for all $i < k$, while $D, O \in Y_k^{k-1}$. Moreover, $\bar{p}(v)$ is weakly increasing in $v$. Since the previous steps removed only second-round prices below $\bar{p}(v_i)$ for $i < k$, $\bar{p}(v_k) \in Y_0^{k-1}$. Hence, the conditions of Lemma A.1 hold. Part (i) removes every remaining second-round price below $\bar{p}(v_k)$, and part (ii) then eliminates type $v_k$'s voluntary delay. This completes the $k$-th step and yields the $k$-th-order deletion of voluntary delay.

Induction therefore yields the $k$-th-order deletion of voluntary delay for every $k = 1, \ldots, N$ when $\epsilon = 0$, and for every $k = 1, \ldots, K^*$ when $\epsilon > 0$. Since $p_1 \in \mathcal{P}$ was arbitrary, the conclusion holds in every $p_1$-continuation game.

### A.3 Maximal Conditional-Dominance Procedure

We now compare the low-to-high deletion chain with the maximal conditional-dominance procedure of Shimoji and Watson (1998).[30] Fix $p_1 \in \mathcal{P}$ as before, and suppress the dependence on $p_1$ in the notation. For any nonempty product set $Y = \prod_{i \in I} Y_i \subseteq S^\epsilon$, let

$$UD_i(Y; \mathcal{A}^\epsilon) := \{s_i \in Y_i : s_i \text{ is not conditionally dominated on } (\mathcal{A}_i^\epsilon, Y)\}$$

and

$$UD(Y; \mathcal{A}^\epsilon) := \prod_{i \in I} UD_i(Y; \mathcal{A}^\epsilon).$$

Define $UD^0(S^\epsilon; \mathcal{A}^\epsilon) := S^\epsilon$ and, recursively,

$$UD^{r+1}(S^\epsilon; \mathcal{A}^\epsilon) := UD\big(UD^r(S^\epsilon; \mathcal{A}^\epsilon); \mathcal{A}^\epsilon\big) \qquad \forall r \geq 0.$$

Thus, at each step, the maximal procedure removes all strategies that are currently conditionally dominated. Let

$$Q(\mathcal{A}^\epsilon) \equiv \prod_{i \in I} Q_i(\mathcal{A}^\epsilon) := \bigcap_{r=0}^{\infty} UD^r(S^\epsilon; \mathcal{A}^\epsilon)$$

denote the set of strategies surviving the maximal procedure. Because the strategy spaces are finite, the maximal deletion procedure stabilizes after finitely many rounds.

**Lemma A.2.** *For any $p_1$-continuation game, $Q_i(\mathcal{A}^\epsilon) \neq \varnothing$ for every $i \in I$.*

*Proof.* Fix any $p_1 \in \mathcal{P}$ and consider the $p_1$-continuation game. Let

$$Y^r := UD^r(S^\epsilon; \mathcal{A}^\epsilon) = \prod_{i \in I} Y_i^r.$$

By construction, $Y_i^0 = S_i^\epsilon \neq \varnothing$ for every $i \in I$. Now consider any integer $r \geq 0$ such that $Y_i^r \neq \varnothing$ for every $i \in I$. We show that $Y_i^{r+1}$ is also nonempty for any $i$.

Write $Y := Y^r = \prod_{i \in I} Y_i$, and let $X^i = X_i^i \times X_{-i}^i$ denote player $i$'s unique restriction in $\mathcal{A}_i^\epsilon$. Recall that $X_i^i = S_i^\epsilon$. If $X_{-i}^i \cap Y_{-i} = \varnothing$, no strategy of player $i$ is conditionally dominated on $(\mathcal{A}_i^\epsilon, Y)$ by definition, so $UD_i(Y; \mathcal{A}^\epsilon) = Y_i \neq \varnothing$. Otherwise, choose a full-support belief $\mu_i \in \Delta(X_{-i}^i \cap Y_{-i})$ and a pure strategy $s_i^* \in Y_i$ that is a best response to $\mu_i$ among all strategies in $Y_i$. Then $s_i^*$ cannot be strictly dominated in $X^i \cap Y$. Indeed, if some $\alpha_i \in \Delta(X_i^i \cap Y_i) = \Delta(Y_i)$ yielded a strictly higher

[30] For finite complete-information extensive games with perfect recall and no moves of Nature, Chen and Micali (2013) show that different conditional-dominance elimination orders may leave different surviving strategy sets, but these sets are equivalent in the sense that suitable player-by-player mappings preserve the terminal node induced by every surviving strategy profile. Their result does not directly apply to our type-agent representation of a Bayesian game with Nature's intervention. Proposition A.1 therefore verifies directly that the maximal procedure eliminates voluntary delay for every buyer type covered by the low-to-high chain.

payoff than $s_i^*$ against every $s_{-i} \in X_{-i}^i \cap Y_{-i}$, it would also yield a strictly higher expected payoff under $\mu_i$, contradicting the optimality of $s_i^*$. Hence, $UD_i(Y; \mathcal{A}^\epsilon) \neq \varnothing$.

It follows by induction that every iteration of the maximal procedure has nonempty strategy components. Since $S^\epsilon$ is finite, the decreasing sequence of survivor sets stabilizes after finitely many rounds. Therefore, $Q_i(\mathcal{A}^\epsilon) \neq \varnothing$ for every $i \in I$. □

**Proposition A.1.** *Suppose Assumptions (A1) and (A2) hold. For every fixed $p_1 \in \mathcal{P}$, the following statements hold.*

*(i) If $\epsilon = 0$, then $D \notin Q_k(\mathcal{A}^0)$ for all $k = 1, 2, \ldots, N$.*

*(ii) If $\epsilon > 0$ and Assumption (A3) also holds, then $D \notin Q_k(\mathcal{A}^\epsilon)$ for all $k = 1, \ldots, K^*$.*

*Proof.* Fix $p_1 \in \mathcal{P}$, and let $Y^r := UD^r(S^\epsilon; \mathcal{A}^\epsilon)$. Since the strategy space is finite, there exists a finite integer $R$ such that $Y^R = Q(\mathcal{A}^\epsilon)$. Throughout this proof, we also write $Q := Q(\mathcal{A}^\epsilon) = \prod_{i \in I} Q_i(\mathcal{A}^\epsilon)$. By Lemma A.2, each $Q_i(\mathcal{A}^\epsilon)$ is nonempty. Since the maximal procedure has stabilized at $Q$, we also have $UD(Q; \mathcal{A}^\epsilon) = Q$. Thus, no strategy in $Q_i(\mathcal{A}^\epsilon)$ is conditionally dominated on $(\mathcal{A}_i^\epsilon, Q)$. Recall that, for either $\epsilon = 0$ or $\epsilon > 0$, each $\mathcal{A}_i^\epsilon$ contains a unique restriction, which we denote by $X^i = X_i^i \times X_{-i}^i$ in this proof. Note that $X_i^i = S_i^\epsilon$ in every case considered here.

*Step I:* Suppose $X_{-i}^i \cap Q_{-i} \neq \varnothing$. For every $s_i \notin Q_i(\mathcal{A}^\epsilon)$, there exists $\alpha_i \in \Delta(Q_i(\mathcal{A}^\epsilon))$ such that

$$U_i^{p_1,\epsilon}(\alpha_i, s_{-i}) > U_i^{p_1,\epsilon}(s_i, s_{-i}) \qquad \forall s_{-i} \in X_{-i}^i \cap Q_{-i}.$$

Suppose otherwise. By the standard strict-dominance/best-response duality (Pearce, 1984, Lemma 3), there exists a belief $\mu_i$ over $X_{-i}^i \cap Q_{-i}$ such that

$$U_i^{p_1,\epsilon}(s_i, \mu_i) \geq U_i^{p_1,\epsilon}(s_i', \mu_i) \qquad \forall s_i' \in Q_i(\mathcal{A}^\epsilon).$$

Let $t_i$ be a best response to $\mu_i$ among all strategies in $S_i^\epsilon$. We claim that $t_i \in Q_i(\mathcal{A}^\epsilon)$. Otherwise, let $r$ be the round in which $t_i$ is eliminated. At that round, some mixed strategy supported on $Y_i^{r-1}$ strictly dominates $t_i$ against every opponent strategy in $X_{-i}^i \cap Y_{-i}^{r-1}$. Since $X_{-i}^i \cap Q_{-i} \subseteq X_{-i}^i \cap Y_{-i}^{r-1}$, the same mixed strategy yields a strictly higher expected payoff than $t_i$ under $\mu_i$, contradicting the optimality of $t_i$. Hence $t_i \in Q_i(\mathcal{A}^\epsilon)$.

Since $t_i \in Q_i(\mathcal{A}^\epsilon)$, the preceding inequality gives $U_i^{p_1,\epsilon}(s_i, \mu_i) \geq U_i^{p_1,\epsilon}(t_i, \mu_i)$. The reverse inequality also follows from the optimality of $t_i$ over $S_i^\epsilon$. Hence $s_i$ is also a best response to $\mu_i$ among all strategies in $S_i^\epsilon$. But $s_i$ was eliminated at some round of the maximal procedure. At that round, a mixed strategy strictly dominated $s_i$ against a set containing the support of $\mu_i$, contradicting the optimality of $s_i$ under $\mu_i$. Therefore, the required mixed strategy $\alpha_i$ exists.

*Step II:* Let $K^\dagger := N$ if $\epsilon = 0$ and $K^\dagger := K^*$ if $\epsilon > 0$. We prove by induction that

$$D \notin Q_k(\mathcal{A}^\epsilon) \qquad \forall k = 1, \dots, K^\dagger.$$

Fix $k \leq K^\dagger$ and suppose $D \notin Q_i(\mathcal{A}^\epsilon)$ for every $i < k$. Suppose, toward a contradiction, that $D \in Q_k(\mathcal{A}^\epsilon)$.

We first show that no second-round price below $\bar{p}(v_k)$ belongs to $Q_0(\mathcal{A}^\epsilon)$. Fix $p \in Q_0(\mathcal{A}^\epsilon)$ such that $p < \bar{p}(v_k)$. Consider any opponent strategy profile consistent with the seller's information-set restriction and $Q$. In other words, take $s_{-0} \in D^0_{-0} \cap Q_{-0}(\mathcal{A}^0)$ if $\epsilon = 0$, and $s_{-0} \in Q_{-0}(\mathcal{A}^\epsilon)$ if $\epsilon > 0$. Let $J(s_{-0}) := \{j \in \{k, \dots, N\} : s_j = D\}$. For every $j \in J(s_{-0})$,

$$u(v_j) \geq u(v_k) > \bar{p}(v_k) > p,$$

so type $v_j$ accepts both $p$ and $\bar{p}(v_k)$ in round 2. Because the voluntary delays of types below $v_k$ have already been eliminated, only types in $J(s_{-0})$ may voluntarily delay. Moreover, when $\epsilon = 0$, $J(s_{-0})$ is nonempty because the seller's information set is reached only after voluntary delay. The seller's payoff gain from replacing $p$ with $\bar{p}(v_k)$ is

$$U_0^{p_1,\epsilon}(\bar{p}(v_k), s_{-0}) - U_0^{p_1,\epsilon}(p, s_{-0}) = \delta\epsilon\big[\Pi(\bar{p}(v_k)) - \Pi(p)\big] + \delta(1-\epsilon)[\bar{p}(v_k) - p] \sum_{j \in J(s_{-0})} q(v_j).$$

If $\epsilon = 0$, the first term is zero and the second term is strictly positive because $J(s_{-0})$ is nonempty. If $\epsilon > 0$, then $p < \bar{p}(v_k) \leq p^*$, so Assumption (A3) implies that the first term is strictly positive, while the second term is weakly positive. Hence, in either case,

$$U_0^{p_1,\epsilon}(\bar{p}(v_k), s_{-0}) > U_0^{p_1,\epsilon}(p, s_{-0})$$

for every surviving opponent strategy profile consistent with the seller's information-set restriction.

Note that $X^0_{-0} \cap Q_{-0} \neq \varnothing$. When $\epsilon = 0$, the contradiction hypothesis $D \in Q_k(\mathcal{A}^0)$ implies $D^0_{-0} \cap Q_{-0}(\mathcal{A}^0) \neq \varnothing$. When $\epsilon > 0$, nonemptiness follows from Lemma A.2. There are two subcases. If $\bar{p}(v_k) \in Q_0(\mathcal{A}^\epsilon)$, let $\alpha_0$ assign probability one to $\bar{p}(v_k)$. If $\bar{p}(v_k) \notin Q_0(\mathcal{A}^\epsilon)$, the observation in Step I provides a mixed strategy $\alpha_0 \in \Delta(Q_0(\mathcal{A}^\epsilon))$ that yields a strictly higher payoff than $\bar{p}(v_k)$ against every surviving opponent strategy profile consistent with the seller's restriction. In either case, $\alpha_0 \in \Delta(Q_0(\mathcal{A}^\epsilon))$ and

$$U_0^{p_1,\epsilon}(\alpha_0, s_{-0}) > U_0^{p_1,\epsilon}(p, s_{-0}) \qquad \forall s_{-0} \in X^0_{-0} \cap Q_{-0}.$$

Since $X^0_0 = S^\epsilon_0$, we have $\alpha_0 \in \Delta(X^0_0 \cap Q_0)$. Thus, $p$ is strictly dominated in $X^0 \cap Q$ and hence conditionally dominated on $(\mathcal{A}^\epsilon_0, Q)$. This contradicts $p \in Q_0(\mathcal{A}^\epsilon)$ and the stability $UD(Q; \mathcal{A}^\epsilon) = Q$.

Therefore,

$$p_2 \geq \bar{p}(v_k) \qquad \forall p_2 \in Q_0(\mathcal{A}^\epsilon).$$

*Step III:* It remains to eliminate type $v_k$'s voluntary delay. By Step II, every $p_2 \in Q_0(\mathcal{A}^\epsilon)$ satisfies $p_2 \geq \bar{p}(v_k)$. For each such $p_2$, let

$$c(v_k, p_2) := \delta \max\{w, v_k - p_2\}$$

denote buyer-agent $k$'s continuation payoff. Then

$$c(v_k, p_2) \leq \delta \max\{w, v_k - \bar{p}(v_k)\} = \delta\big[w + u(v_k) - \bar{p}(v_k)\big] \leq \delta(w + \Delta) < w,$$

where the equality follows from $\bar{p}(v_k) < u(v_k) = v_k - w$, and the last inequality follows from Assumption (A2). Choosing $D$ yields $c(v_k, p_2)$, whereas choosing $O$ yields $(1-\epsilon)w + \epsilon c(v_k, p_2)$. Hence,

$$U_k^{p_1,\epsilon}(p_2, O) - U_k^{p_1,\epsilon}(p_2, D) = (1-\epsilon)\big[w - c(v_k, p_2)\big] > 0 \qquad \forall p_2 \in Q_0(\mathcal{A}^\epsilon),$$

where the strict inequality uses $\epsilon < 1$. Thus, $O$ yields a strictly higher payoff than $D$ against every surviving seller price.

Let $b_k(p_1) := A$ if $u(v_k) > p_1$ and $b_k(p_1) := O$ otherwise. By construction, $b_k(p_1)$ yields a payoff weakly greater than $O$ against every seller price. Hence, $b_k(p_1)$ also yields a strictly higher payoff than $D$ against every surviving seller price. By Lemma A.2, $X_{-k}^k \cap Q_{-k} = Q_{-k} \neq \varnothing$. There are two subcases. If $b_k(p_1) \in Q_k(\mathcal{A}^\epsilon)$, let $\alpha_k$ assign probability one to $b_k(p_1)$. If $b_k(p_1) \notin Q_k(\mathcal{A}^\epsilon)$, the observation in Step I provides a mixed strategy $\alpha_k \in \Delta(Q_k(\mathcal{A}^\epsilon))$ that strictly dominates $b_k(p_1)$ against every surviving opponent strategy profile. In either case,

$$U_k^{p_1,\epsilon}(\alpha_k, s_{-k}) > U_k^{p_1,\epsilon}(D, s_{-k}) \qquad \forall s_{-k} \in X_{-k}^k \cap Q_{-k}.$$

Since $X_k^k = S_k^\epsilon$, we have $\alpha_k \in \Delta(X_k^k \cap Q_k)$. Thus, $D$ is strictly dominated in $X^k \cap Q$ and hence conditionally dominated on $(\mathcal{A}_k^\epsilon, Q)$. This contradicts $D \in Q_k(\mathcal{A}^\epsilon)$ and the stability $UD(Q; \mathcal{A}^\epsilon) = Q$. Therefore, $D \notin Q_k(\mathcal{A}^\epsilon)$. $\square$

### A.4 Proof of Proposition 2

Let $\hat{q} = \hat{q}^{\sigma_B}$ for some buyer behavioral strategy profile $\sigma_B$ surviving the $k$-th-order deletion. For each $p_1 \in \mathcal{P}$, define

$$Z_k^{\sigma_B}(p_1) := \epsilon + (1-\epsilon) \sum_{\ell=k+1}^{N} q(v_\ell)\sigma_{v_\ell}(D|p_1).$$

For every $i = 1, \ldots, k$, equation (2.5) implies

$$R_i(p_1) = \frac{\epsilon}{Z_k^{\sigma_B}(p_1)} \le 1,$$

where the inequality follows from $Z_k^{\sigma_B}(p_1) \ge \epsilon$. Next, we show that $R_1(p_1) = \cdots = R_k(p_1)$. For every $j > k$, equation (2.4) gives

$$R_j(p_1) = \frac{\epsilon + (1-\epsilon)\sigma_{v_j}(D|p_1)}{Z_k^{\sigma_B}(p_1)} \ge \frac{\epsilon}{Z_k^{\sigma_B}(p_1)} = R_k(p_1).$$

Therefore, $\hat{q}$ satisfies $k$-th-order positive selection, and (2.6) also holds.

### A.5 Proof of Proposition 3

Suppose $\epsilon = 0$. An argument essentially identical to the low-to-high deletion argument in the proof of Proposition 1 shows that no voluntary delay occurs after any first-round price in any PBE. Indeed, if $v_k$ were the lowest type that delays with positive probability, then, conditional on continuation, the seller would assign zero probability to all types below $v_k$ and would never optimally offer a second-round price below $\bar{p}(v_k)$. Type $v_k$'s payoff from delaying would therefore be at most $\delta[v_k - \bar{p}(v_k)] \le \delta(w + \Delta) < w$, contradicting the payoff $w$ from exercising the outside option immediately. Hence, after any first-round price $p_1$, each buyer type accepts if and only if $u(v) > p_1$ and otherwise exercises the outside option. The seller's expected payoff from offering $p_1$ is therefore $\Pi(p_1)$. Since $p^*$ is the unique maximizer of $\Pi$, the seller offers $p^*$ in the first round.

Next, let $\epsilon^n \downarrow 0$ with $\epsilon^n > 0$ for every $n$, and for each $n$ let $(\sigma_B^n, \sigma_S^n, \hat{q}^n)$ be a PBE assessment of the game with perturbation $\epsilon^n$, where $\sigma_B^n$ and $\sigma_S^n$ denote the buyer's and seller's behavioral strategies. Because the seller does not observe the buyer's first-round response, the continuation payoff following Nature's intervention is the same under $A$ and $O$. The difference between the payoffs from $A$ and $O$ in response to $p_1$ in round 1 is $(1-\epsilon^n)(u(v) - p_1)$. Therefore, when $u(v) > p_1$, $O$ is strictly inferior to $A$, whereas when $u(v) < p_1$, $A$ is strictly inferior to $O$. It therefore suffices to establish

$$\lim_{n\to\infty} \sigma_v^n(D|p_1) = 0 \qquad \forall p_1 \in \mathcal{P},\ \forall v \in V. \tag{A.1}$$

Granting (A.1), the seller's expected payoff from offering $p_1$ in $(\sigma_B^n, \sigma_S^n, \hat{q}^n)$ is $\Pi(p_1) + o_n(p_1)$, where $o_n(p_1) \to 0$ uniformly over $p_1$ because $\mathcal{P}$ is finite. Since $p^*$ is the unique maximizer of $\Pi$ over the finite set $\mathcal{P}$, it follows that the seller's equilibrium first-round offer equals $p^*$ for all sufficiently large $n$.

We now prove (A.1). Pick $p_1 \in \mathcal{P}$ arbitrarily, and suppose for contradiction that there is a buyer type $v \in V$ such that $\limsup_{n\to\infty} \sigma_v^n(D|p_1) > 0$. Let $\hat{v}$ be the lowest one among all such buyer types; that is, $\lim_{n\to\infty} \sigma_v^n(D|p_1) = 0$ for any $v < \hat{v}$. Passing to a subsequence, we may assume that $\sigma_{\hat{v}}^n(D|p_1)$

converges to a strictly positive limit and that $\sigma_v^n(D|p_1)$ converges for every $v \in V$. Suppose the negotiation reaches round 2 after the first-round price $p_1$, and the seller offers $p_2 = p_2' < \overline{p}(\hat{v})$. Since $p_2'$ is accepted if and only if $u(v) > p_2'$, the seller's undiscounted second-round expected revenue is

$$R_2^n(p_2'|p_1) := p_2' \sum_{v:u(v)>p_2'} \hat{q}^n(v|p_1).$$

Similarly, the seller's payoff from charging $p_2 = \overline{p}(\hat{v})$ in the second round is

$$R_2^n(\overline{p}(\hat{v})|p_1) := \overline{p}(\hat{v}) \sum_{v:u(v)>\overline{p}(\hat{v})} \hat{q}^n(v|p_1).$$

The difference between these two payoffs is

$$\begin{aligned} R_2^n(\overline{p}(\hat{v})|p_1) - R_2^n(p_2'|p_1) &= (\overline{p}(\hat{v}) - p_2') \sum_{u(v)>\overline{p}(\hat{v})} \hat{q}^n(v|p_1) - p_2' \sum_{p_2'<u(v)<\overline{p}(\hat{v})} \hat{q}^n(v|p_1) \\ &\geq (\overline{p}(\hat{v}) - p_2')\hat{q}^n(\hat{v}|p_1) - p_2' \sum_{p_2'<u(v)<\overline{p}(\hat{v})} \hat{q}^n(v|p_1) \\ &= \hat{q}^n(\hat{v}|p_1)\left[\overline{p}(\hat{v}) - p_2' - p_2' \sum_{p_2'<u(v)<\overline{p}(\hat{v})} \frac{\hat{q}^n(v|p_1)}{\hat{q}^n(\hat{v}|p_1)}\right], \end{aligned}$$

where the weak inequality follows because $u(\hat{v}) > \overline{p}(\hat{v})$, so type $\hat{v}$ is included in the first sum. By Bayes' rule,

$$\lim_{n\to\infty} \frac{\hat{q}^n(v|p_1)}{\hat{q}^n(\hat{v}|p_1)} = \lim_{n\to\infty} \frac{q(v)[\epsilon^n + (1-\epsilon^n)\sigma_v^n(D|p_1)]}{q(\hat{v})[\epsilon^n + (1-\epsilon^n)\sigma_{\hat{v}}^n(D|p_1)]} = \frac{q(v)\cdot 0}{q(\hat{v})\lim_{n\to\infty}\sigma_{\hat{v}}^n(D|p_1)} = 0$$

for any $v$ such that $u(v) < \overline{p}(\hat{v})$. On the other hand,

$$\begin{aligned} \lim_{n\to\infty} \hat{q}^n(\hat{v}|p_1) &= \lim_{n\to\infty} \frac{q(\hat{v})[\epsilon^n + (1-\epsilon^n)\sigma_{\hat{v}}^n(D|p_1)]}{\epsilon^n + (1-\epsilon^n)\sum_v q(v)\sigma_v^n(D|p_1)} \\ &\geq \lim_{n\to\infty} q(\hat{v})\left(\epsilon^n + (1-\epsilon^n)\sigma_{\hat{v}}^n(D|p_1)\right) \\ &= q(\hat{v}) \lim_{n\to\infty} \sigma_{\hat{v}}^n(D|p_1) \\ &> 0, \end{aligned}$$

where the weak inequality holds because the denominator $\epsilon^n + (1-\epsilon^n)\sum_v q(v)\sigma_v^n(D|p_1)$ is always weakly less than 1. Thus,

$$\lim_{n\to\infty}\left(R_2^n(\overline{p}(\hat{v})|p_1) - R_2^n(p_2'|p_1)\right) \geq \lim_{n\to\infty} \hat{q}^n(\hat{v}|p_1)(\overline{p}(\hat{v}) - p_2') > 0 \quad \forall p_2' < \overline{p}(\hat{v}).$$

Since $\mathcal{P}$ is finite, it follows that, for all sufficiently large $n$,

$$R_2^n(\overline{p}(\hat{v})|p_1) > R_2^n(p_2'|p_1) \qquad \forall p_2' < \overline{p}(\hat{v}).$$

Thus, for all sufficiently large $n$, the seller never offers a second-round price below $\overline{p}(\hat{v})$ following continuation after $p_1$. Let $c_n$ denote type $\hat{v}$'s equilibrium payoff conditional on reaching round 2 after $p_1$. Then

$$c_n \leq \delta \max\{w, \hat{v} - \bar{p}(\hat{v})\} = \delta\big[w + u(\hat{v}) - \bar{p}(\hat{v})\big] \leq \delta(w + \Delta) < w,$$

where the equality follows from $\bar{p}(\hat{v}) < u(\hat{v}) = \hat{v} - w$, and the last inequality follows from Assumption (A2). If type $\hat{v}$ voluntarily delays after $p_1$, she obtains $c_n$. If she instead exercises the outside option, she obtains $w$ when Nature does not intervene and the same continuation payoff $c_n$ when Nature forces continuation. Hence, the payoff gain from exercising the outside option rather than delaying is $(1 - \epsilon^n)(w - c_n) > 0$. Thus, voluntary delay cannot be optimal for type $\hat{v}$, contradicting $\limsup_{n\to\infty} \sigma_{\hat{v}}^n(D|p_1) > 0$.

It remains to establish the uniform threshold. Suppose no such $\bar{\epsilon}$ exists. Then, for every $m \in \mathbb{N}$, there are $\epsilon^m \in [0, 1/m)$ and a PBE of the game with perturbation $\epsilon^m$ in which the seller does not offer $p^*$ with probability one in the first round. By part (i), $\epsilon^m > 0$ for every $m$. Passing to a decreasing subsequence if necessary, we may assume that $\epsilon^m \downarrow 0$. The sequence $(\epsilon^m)_{m=1}^{\infty}$ and the associated PBE therefore contradict the conclusion of the preceding paragraph. Hence $\bar{\epsilon}$ exists.

# B Two-Round Bargaining Game with No Outside Option

This appendix derives the equilibrium of the two-round bargaining game used in the supplementary experiment reported in Section 5.1, in which the buyer has no outside option. We represent the absence of an outside option by setting $w = -\infty$ and consider $v_L = 20$, $v_M \in \{40, 190, 370\}$, $v_H = 450$, $\Delta = 10$, $\delta = 0.8$, and $\epsilon = 10^{-3}$, and maintain the uniform prior $q_L = q_M = q_H = 1/3$. The buyer valuations are chosen so that each type's gain from trade coincides with its net valuation in the corresponding environment with $w = 50$. The associated prices are $p_L = v_L - \Delta = 10$, $p_M = v_M - \Delta \in \{30, 180, 360\}$, $p_H = v_H - \Delta = 440$. The static monopoly price is

$$p^\star = \arg\max_{p \in \{p_L, p_M, p_H\}} p \sum_{k: v_k > p} q_k = \begin{cases} p_H, & \text{if } v_M \in \{40, 190\}, \\ p_M, & \text{if } v_M = 370. \end{cases}$$

## B.1 Preliminary Observations

Note that, for all possible values of $v_M$,

$$p_M(x + q_M) > p_L(x + q_M + q_L) \quad \forall x \in [0, 1/3]. \tag{B.1}$$

Thus, if the negotiation reaches the second round following $p_1 = p_H$, the seller will never offer $p_2 = p_L$ in the second round. Also,

$$(1-\epsilon)(v_H - p_H) + \epsilon\delta(v_H - p_M) < \delta(v_H - p_M) \quad \forall v_M \in \{40, 190, 370\}, \tag{B.2}$$

which implies that if the high type expects $p_2 = p_M$, she will never accept $p_1 = p_H$. Similarly,

$$(1-\epsilon)(v_M - p_M) + \epsilon\delta(v_M - p_L) < \delta(v_M - p_L) \quad \forall v_M \in \{40, 190, 370\}, \tag{B.3}$$

which implies that if the middle type expects $p_2 = p_L$, she will never accept $p_1 = p_M$.

## B.2 Analysis of Continuation Games

**Case I: Continuation Game After $p_1 = p_H$** First, focus on the case $v_M \in \{40, 190\}$. Suppose that the seller offers $p_1 = p_H$ in the first round. The middle-type and low-type buyers would reject this price, as their values are strictly less than $p_H$. We claim that the high type would randomize.

- Suppose, for contradiction, that the high type accepts this offer with probability one in equilibrium. Then, the seller's posterior beliefs in the second round would be

  $$\hat{q}_H = \frac{\epsilon}{2+\epsilon} \approx 0 \quad \text{and} \quad \hat{q}_M = \hat{q}_L = \frac{1}{2+\epsilon} \approx \frac{1}{2}.$$

  Since $p_M(\hat{q}_M + \hat{q}_H) > p_H \hat{q}_H$ for $v_M \in \{40, 190\}$, the seller never finds it optimal to offer $p_2 = p_H$. By (B.1), the seller will therefore offer $p_2 = p_M$ in the second round. Then, by (B.2), the high type would find it profitable to deviate and reject $p_1 = p_H$.

- Next, suppose that the high type rejects $p_1 = p_H$ with probability one in equilibrium. Then, the seller will retain his prior beliefs in the second round and offer $p_2 = p^\star = p_H$. In other words, the seller will offer the same price in the second round. Thus, the high type's rejection of $p_1 = p_H$ in the first round cannot be justified.

Let $x_H$ denote the high type's probability of accepting $p_1 = p_H$. The seller's equilibrium posterior beliefs are then

$$\begin{aligned}
\hat{q}_H(p_H) &= \frac{q_H(x_H\epsilon + 1 - x_H)}{q_L + q_M + q_H(x_H\epsilon + 1 - x_H)}, \\
\hat{q}_M(p_H) &= \frac{q_M}{q_L + q_M + q_H(x_H\epsilon + 1 - x_H)}, \\
\hat{q}_L(p_H) &= \frac{q_L}{q_L + q_M + q_H(x_H\epsilon + 1 - x_H)}.
\end{aligned} \tag{B.4}$$

The high type's randomization is justified only if the seller also randomizes in the second round. By (B.1), the only possibility is that the seller randomizes over $p_2 = p_H$ and $p_2 = p_M$ with probabilities

$y_H$ and $1 - y_H$, respectively. To justify the seller's randomization, the seller must be indifferent between the two offers:

$$p_M[\hat{p}_H(p_H) + \hat{p}_M(p_H)] = p_H \hat{p}_M(p_H) \iff x_H = \frac{1}{1-\epsilon}\left[1 - \frac{q_M}{q_H}\frac{p_M}{p_H - p_M}\right].$$

Here, $x_H$ indeed lies in $(0,1)$ when $v_M \in \{40, 190\}$. In turn, the high type is willing to mix only if the seller assigns probabilities $y_H$ and $1 - y_H$ to $p_2 = p_H$ and $p_2 = p_M$, respectively (recall that, by (B.1), the seller will never offer $p_2 = p_L$), where

$$v_H - p_H = \delta y_H (v_H - p_H) + \delta(1 - y_H)(v_H - p_M) \iff y_H = 1 - \frac{(1-\delta)(v_H - p_H)}{\delta(p_H - p_M)}.$$

One can easily verify that $y_H$ indeed lies in $(0,1)$ when $v_M \in \{40, 190\}$.

**Observation 1.** *Suppose $v_M \in \{40, 190\}$. Suppose that the seller offers $p_1 = p_H$ in the first round. In equilibrium, the high type accepts this offer with probability $x_H$, and then, in the second round, the seller randomizes over $p_H$ and $p_M$ with probabilities $y_H$ and $1 - y_H$, respectively, where*

$$x_H = \frac{1}{1-\epsilon}\left[1 - \frac{q_M}{q_H}\frac{p_M}{p_H - p_M}\right] \quad \text{and} \quad y_H = 1 - \frac{(1-\delta)(v_H - p_H)}{\delta(p_H - p_M)}.$$

*Thus, the seller's profit from offering $p_1 = p_H$ is $\pi_H = p_H q_H [x_H(1-\epsilon) + \delta x_H \epsilon + \delta(1 - x_H)]$. The seller's posterior beliefs after continuation following $p_1 = p_H$ are*

$$\hat{q}_H(p_H) = \frac{q_H[x_H\epsilon + 1 - x_H]}{q_L + q_M + q_H[x_H\epsilon + 1 - x_H]} = \frac{x_H\epsilon + 1 - x_H}{2 + x_H\epsilon + 1 - x_H} < q_H = \frac{1}{3},$$
$$\hat{q}_k(p_H) = \frac{q_k}{q_L + q_M + q_H[x_H\epsilon + 1 - x_H]} = \frac{1}{2 + [x_H\epsilon + 1 - x_H]} > q_k = \frac{1}{3} \quad k = L, M.$$

*Note that negative selection occurs in this case: relative to the prior, the posterior shifts probability mass toward lower types and away from higher types.*

What if $v_M = 370$? The high type must reject $p_1 = p_H$ with positive probability, say $1 - x_H > 0$. Suppose, for contradiction, that the high type accepts $p_1 = p_H$ with probability one. Note that the seller's posterior beliefs in the second round are still given by (B.4). With $x_H = 1$, the seller's posterior beliefs in the second round would be

$$\hat{q}_H = \frac{\epsilon}{2+\epsilon} \approx 0 \quad \text{and} \quad \hat{q}_M = \hat{q}_L = \frac{1}{2+\epsilon} \approx \frac{1}{2}.$$

By (B.1) and since $p_M(\hat{q}_M + \hat{q}_H) > p_H \hat{q}_H$, the seller would offer $p_2 = p_M$ in the second round. Then, by (B.2), the high type would profitably deviate by rejecting $p_1 = p_H$.

We claim that $x_H$ must be zero. Suppose, for contradiction, that $x_H \in (0,1)$. Again by (B.1), the seller would not offer $p_2 = p_L$ in the second round. Also, by (B.2), the high type would reject

$p_1 = p_H$ with probability one if the seller is expected to offer $p_2 = p_M$ with probability one. On the other hand, if the seller is expected to offer $p_2 = p_H$ with probability one, the high type would accept $p_1 = p_H$ with probability one. Hence, the only remaining possibility is that the seller randomizes over $p_2 = p_H$ and $p_2 = p_M$ in the second round, which requires the seller to be indifferent between $p_2 = p_H$ and $p_2 = p_M$ under the posterior beliefs:

$$p_M[q_M + q_H(x_H\epsilon + 1 - x_H)] = p_H[q_H(x_H\epsilon + 1 - x_H)] \iff x_H = \frac{1}{1-\epsilon}\left[1 - \frac{q_M}{q_H}\frac{p_M}{p_H - p_M}\right].$$

But when $v_M = 370$, this expression for $x_H$ is negative, which is impossible.

Thus, if $v_M = 370$, we must have $x_H = 0$; that is, the high type rejects $p_1 = p_H$ with probability one. In the second round, the seller's posterior beliefs coincide with the prior, and thus the seller will offer $p_2 = p^\star = p_M$. Thus, with $v_M = 370$, the seller's profit from $p_1 = p_H$ is

$$\pi_H = \delta p_M(q_H + q_M) = \frac{2}{3}\delta p_M.$$

**Observation 2.** *Suppose $v_M = 370$. Suppose that the seller offers $p_1 = p_H$ in the first round. In equilibrium, the high type accepts this offer with probability zero, i.e., $x_H = 0$. In the second round, the seller offers $p_2 = p^\star = p_M$, which is accepted by the high and middle types. Thus, the seller's profit from offering $p_1 = p_H$ is $\pi_H = \delta p_M(q_H + q_M)$. The seller's posterior beliefs after continuation following $p_1 = p_H$ coincide with the prior.*

**Case II: Continuation Game After $p_1 = p_M$** Consider the case in which the seller offers $p_1 = p_M$. We claim that the middle type randomizes.

- Suppose that the middle type accepts $p_1 = p_M$ with probability one, which implies that the high type also accepts $p_1 = p_M$. Then, the seller's posterior satisfies $\hat{q}_L(p_M) = 1/(1+2\epsilon) \approx 1$. Moreover, $p_L(q_L + \epsilon q_M + \epsilon q_H) > \max\{p_M\epsilon(q_M + q_H), p_H\epsilon q_H\}$ for all three values of $v_M$, so the seller offers $p_2 = p_L$ in the second round. But then, by (B.3), the middle type would find it suboptimal to accept $p_1 = p_M$. Contradiction.

- Next, suppose that the middle type rejects $p_1 = p_M$ with probability one. By (B.1), however, the second-round offer $p_2 = p_L$ is strictly dominated by $p_2 = p_M$ for the seller. Hence, every second-round best response has support contained in $\{p_M, p_H\}$. The middle type's payoff from voluntarily delaying is therefore at most $\delta(v_M - p_M)$, whereas accepting $p_1 = p_M$ yields at least $(1-\epsilon)(v_M - p_M)$. Since $1 - \epsilon > \delta$, the middle type strictly prefers accepting $p_1 = p_M$ to voluntarily delaying. Contradiction.

By the single-crossing property, if the middle type accepts $p_1 = p_M$ with positive probability, the high type accepts $p_1 = p_M$ with probability one. Let $x_M$ denote the middle type's probability of accepting $p_1 = p_M$. The seller's equilibrium posterior beliefs are then

$$\hat{q}_H(p_M) = \frac{q_H\epsilon}{q_L + q_M(x_M\epsilon + 1 - x_M) + q_H\epsilon},$$
$$\hat{q}_M(p_M) = \frac{q_M(x_M\epsilon + 1 - x_M)}{q_L + q_M(x_M\epsilon + 1 - x_M) + q_H\epsilon},$$
$$\hat{q}_L(p_M) = \frac{q_L}{q_L + q_M(x_M\epsilon + 1 - x_M) + q_H\epsilon}.$$

For example, when $v_M = 190$, $(\hat{q}_L(p_M), \hat{q}_M(p_M), \hat{q}_H(p_M)) \approx (0.944, 0.055, 0.001)$. Also, note that $\hat{q}_H(p_M) \approx 0$.

The middle type's randomization is justified only if the seller also randomizes in the second round. Because $p_H q_H \epsilon < p_L q_L$, the seller finds it suboptimal to offer $p_2 = p_H$ after continuation following $p_1 = p_M$. Suppose that the seller offers $p_2 = p_M$ and $p_2 = p_L$ with probabilities $y_M$ and $1 - y_M$, respectively. To ensure that the seller is indifferent between the two prices, we must have

$$p_L = p_M \frac{q_M(x_M\epsilon + 1 - x_M) + q_H\epsilon}{q_L + q_M(x_M\epsilon + 1 - x_M) + q_H\epsilon} \iff x_M = \frac{1}{1-\epsilon}\left(1 - \frac{p_L q_L - q_H\epsilon(p_M - p_L)}{q_M(p_M - p_L)}\right).$$

Here, $x_M$ indeed lies in $(0,1)$ in all three cases $v_M \in \{40, 190, 370\}$. In turn, the middle type's randomization $x_M \in (0,1)$ is justified only if the seller also randomizes over $p_2 = p_M$ and $p_2 = p_L$ with probabilities $y_M$ and $1 - y_M$, respectively, where

$$v_M - p_M = \delta y_M(v_M - p_M) + \delta(1 - y_M)(v_M - p_L) \iff y_M = 1 - \frac{(1-\delta)(v_M - p_M)}{\delta(p_M - p_L)}.$$

One can easily verify that $y_M$ indeed lies in $(0,1)$ in all three cases $v_M \in \{40, 190, 370\}$.

**Observation 3.** *Suppose that the seller offers $p_1 = p_M$. In all three cases $v_M \in \{40, 190, 370\}$, the middle type accepts this offer with probability $x_M$, and then, in the second round, the seller randomizes over $p_2 = p_M$ and $p_2 = p_L$ with probabilities $y_M$ and $1 - y_M$, respectively, where*

$$x_M = \frac{1}{1-\epsilon}\left(1 - \frac{p_L q_L - q_H\epsilon(p_M - p_L)}{q_M(p_M - p_L)}\right) \in (0,1) \quad \textit{and} \quad y_M = 1 - \frac{(1-\delta)(v_M - p_M)}{\delta(p_M - p_L)} \in (0,1).$$

*Thus, the seller's profit from offering $p_1 = p_M$ is $\pi_M = p_M[q_H + q_M x_M](1-\epsilon) + \delta[q_H\epsilon + q_M x_M\epsilon + q_M(1 - x_M)]p_M$. The seller's posterior beliefs after continuation following $p_1 = p_M$ are*

$$\hat{q}_H(p_M) = \frac{q_H\epsilon}{q_L + q_M(x_M\epsilon + 1 - x_M) + q_H\epsilon},$$
$$\hat{q}_M(p_M) = \frac{q_M(x_M\epsilon + 1 - x_M)}{q_L + q_M(x_M\epsilon + 1 - x_M) + q_H\epsilon},$$
$$\hat{q}_L(p_M) = \frac{q_L}{q_L + q_M(x_M\epsilon + 1 - x_M) + q_H\epsilon},$$

*where $\hat{q}_H(p_M) \approx 0$ and $\hat{q}_M(p_M) < \frac{1}{3} < \hat{q}_L(p_M)$. Negative selection occurs in this case: relative to the prior, the posterior shifts probability mass toward lower types and away from higher types.*

**Case III: Continuation Game After $p_1 = p_L$** All buyer types accept $p_1 = p_L$. If the negotiation nevertheless reaches the second round due to Nature's intervention, the seller's posterior coincides with the prior, so the seller offers $p_2 = p^\star$. The seller's expected payoff from offering $p_1 = p_L$ is

$$\pi_L = p_L(1-\epsilon) + \delta\epsilon p^\star \sum_{k: v_k > p^\star} q_k \approx p_L,$$

where $p^\star = p_M$ if $v_M = 370$ and $p^\star = p_H$ otherwise.

## B.3 Equilibrium

The continuation equilibrium following each possible first-round offer $p_1$ has been characterized above. It remains to determine the seller's first-round offer. We therefore compare the three corresponding expected profits, $\pi_H$, $\pi_M$, and $\pi_L$.

- If $v_M \in \{40, 190\}$, $\pi_H > \pi_M > \pi_L$. In equilibrium, the seller offers $p_1 = p_H$. The equilibrium play on the path then follows Observation 1.
- If $v_M = 370$, $\pi_M > \pi_H > \pi_L$. In equilibrium, the seller offers $p_1 = p_M$. The equilibrium play on the path then follows Observation 3.

Table 6: Equilibrium predictions across candidate values of $v_M$ (no outside option)

| $v_M$ | On the equilibrium path | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $p_1$ | Mixing buyer type | Pr(Reject) | $p_2$ after continuation | Seller's equilibrium profit | Full commitment profit | Gap | On-path posterior |
| 40 | 440 | High | 0.0722 | (440, .994; 30, .006) | 144.520 | 146.637 | 2.117 | (.482, .482, .035) |
| 190 | 440 | High | 0.6920 | (440, .990; 180, .010) | 126.359 | 146.637 | 20.278 | (.371, .371, .257) |
| 370 | 360 | Middle | 0.0266 | (360, .993; 10, .007) | 239.314 | 239.952 | 0.638 | (.972, .027, .001) |

*Notes:* "Mixing buyer type" denotes the buyer type that randomizes between accepting and rejecting the equilibrium first-round offer. Pr(Reject) is this buyer type's equilibrium probability of rejection. Entries in the $p_2$ column are reported as (price, probability) pairs. "Full commitment profit" is the seller's profit from committing to $p^\star$ in both rounds, accounting for Nature's intervention and discounting. "Gap" is the difference between the full-commitment profit and the equilibrium profit. The on-path posterior is reported in the order $(\hat{q}_L, \hat{q}_M, \hat{q}_H)$ after continuation following the equilibrium first-round offer.

# C Supplementary Experiments

This appendix reports the experimental design and key quantitative results for each of the four supplementary experiments summarized in Section 5. Full instructions, additional figures, and disaggregated data are available in the Online Appendix.

## C.1 No-Outside-Option Experiment (supports Section 5.1)

The treatment uses parameter values $v_L = 20$, $v_M = 190$, $v_H = 450$, $\Delta = 10$, $\delta = 0.8$, $\epsilon = 10^{-3}$, and no outside option. These values are calibrated to match the gain from trading in the main treatment M240. The equilibrium predictions are derived in Appendix B and summarized for $v_M = 190$ as follows: in the first round, the seller offers $p_1 = p_H = 440$, and the high-type buyer accepts with probability 0.308. After rejection, the seller's on-equilibrium-path posterior is $(\hat{q}(v_L), \hat{q}(v_M), \hat{q}(v_H)) = (0.371, 0.371, 0.257)$, exhibiting negative selection.

We conducted 5 sessions with 86 participants. The average first-round offer is 305.19, comparable to 295.86 in M240. $p_H$ was offered in 48.60% of first-round observations. After rejection, 97.8% of sellers either decreased or maintained the second-round price. Table 7 reports a summary in the comparable format of Table 4.

| | M240 | NoOutside |
|---|---|---|
| Avg.Offer (Theory) | 295.86 (440) | 305.19 (440) |
| %Delay $v_L$ (Theory) | 35 (0) | 100 (100) |
| %Delay $v_M$ (Theory) | 60 (0) | 84.1 (100) |
| %Delay $v_H$ (Theory) | 50 (0) | 47.9 (69.2) |
| Avg.SellerPayoffs (Theory) | 87.94 (146.67) | 90.53 (126.359) |
| Avg.BuyerPayoffs (Theory) | 119.20 (53.33) | 76.21 (3.36) |

Table 7: Comparisons between M240 and NoOutside

## C.2 No $p_L$ Experiment (supports Section 5.2)

The treatment retains all parameter values from the main experiment except that the seller's price set is restricted to $\{p_M, p_H\}$, removing the lowest price $p_L$. Under this restriction, the low-type buyer cannot afford any feasible second-round price, making the outside option the unique rational choice for the low type rather than merely a strategically dominant one.

We conducted 5 sessions with 90 participants. 88% of low-type buyers (125 out of 142) took the outside option in the first round, and only 2% of high-type buyers (3 out of 154) did so. Because $p_L$ is unavailable, many bargaining matches concluded in the first round: 21.6% of first-round bargaining (97 out of 450 observations) ended in rejection, yielding 74 posterior belief observations. Despite the elimination of $p_L$, sellers continue to assign substantial weight to the low type in their reported posterior beliefs, as displayed in Figure 7 of Section 5.2.

### C.3 Other-regarding preference and the ARP method (supports Section 5.3)

To examine whether other-regarding preferences drive our main results, we conducted three additional sessions, one per main treatment ($v_M \in \{90, 240, 420\}$), incorporating the Amended Random Payment (ARP) method of Lim and Xiong (2021). Under ARP, each participant's final payment is determined by an independently drawn random match. Participants are informed whether the current match determines the paired participant's payment but not which match determines their own. Crucially, in matches that do not determine the paired participant's payment (90% of observations), decisions cannot affect the other participant's earnings, neutralizing other-regarding motives.

We had 12, 14, and 18 participants in M90, M240, and M420, respectively. Average payment was HKD 131.6 (≈ USD 17) including a HKD 40 show-up fee. Table 8 reports the analog of Table 4 for the ARP sessions. Rejection rates remain substantially above the theoretical prediction of zero. In particular, 73% of middle-type buyers reject $p_M$ in M420 under ARP, compared to 74% in the main experiment. The reported posterior beliefs (Figure 8 in Section 5.3) cluster outside the region consistent with first-order positive selection, qualitatively replicating the patterns in Figure 2 of the main experiment.

| | $M90$ | $M240$ | $M420$ |
|---|---|---|---|
| Avg.Offer (Theory) | 333.33 (440) | 306.83 (440) | 353.46 (360) |
| %Delay_$v_L$ (Theory) | 12 (0) | 28 (0) | 39 (0) |
| %Delay_$v_M$ (Theory) | 41 (0) | 53 (0) | 73 (0) |
| %Delay_$v_H$ (Theory) | 60 (0) | 65 (0) | 68 (0) |
| Avg.SellerPayoffs (Theory) | 66.11 (146.67) | 65.24 (146.67) | 101.60 (240) |
| Avg.BuyerPayoffs (Theory) | 108.33 (53.33) | 142.22 (53.33) | 125.43 (83.33) |

Table 8: Summary of Experimental Findings: ARP

Avg.Offer is the average of the first-round price offers. %Delay_$v_i$ is the percentage that a buyer with type $v_i$ voluntarily delays the price offer. Avg.SellerPayoffs and Avg.BuyerPayoffs are the average payoffs of the sellers and the buyers, respectively. Theoretically predicted values are in parentheses.

Compared to the data from the main treatments, the rejection rates are generally lower. However, these rejection rates still remain substantially higher than the theoretical predictions. Overall, the results do not provide evidence that other-regarding preferences qualitatively alter the observed behavior in our experiment.

The posterior beliefs of sellers reported in the three ARP sessions are illustrated in Figure 8. Consistent with the key finding from the belief data in the main treatments, the vast majority of observations lie outside the shaded area. This indicates that the overwhelming majority of sellers fail to exhibit first-order positive selection. Consequently, we can conclude that the main results regarding the failure of positive selection are not an artifact primarily driven by other-regarding preference considerations.

### C.4 Take-it-or-Leave-it Bargaining (supports Section 5.4)

To examine whether the failure of positive selection reflects a more basic limitation in sellers' pricing competence, and to address the related concern that sellers may underprice in anticipation of fairness-motivated rejections, we conducted four additional sessions of an ultimatum bargaining experiment with a privately-informed buyer and an outside option worth 50.

Each session consisted of 18 participants (9 buyers and 9 sellers), with roles fixed across eight randomly rematched matches. The buyer's value was drawn uniformly from $[50, 400]$. The seller made a single take-it-or-leave-it (TIOLI) offer, and the buyer chose between accepting and exercising the outside option. Under these parameters, the offer that maximizes the seller's expected profit is 175.

The average TIOLI offer is 166.92, which is not statistically distinguishable from 175 ($p = 0.391$, two-sided $t$-test). Figure 9 in Section 5.4 plots the buyer's value against the seller's offer; acceptance and rejection are well separated by the buyer's indifference line between accepting and exercising the outside option. Buyers accept inequitable offers when accepting is in their material self-interest, indicating that sellers in the main experiment have no reason to anticipate widespread fairness-motivated rejections.

## D Sample Experimental Instructions

**INSTRUCTIONS (For $v_M = 240$)**

Welcome to the experiment. Please read these instructions carefully. There will be a quiz around the end of the instructions, to make sure you understand this experiment. The payment you will receive from this experiment depends on your decisions.

**Your Role and Match**

At the beginning of the experiment, one-half of the participants will be randomly assigned to the role of a **seller** and the other half the role of a **buyer**. Your role will remain fixed throughout the experiment.

The experiment consists of 10 **matches**. At the beginning of each match, one seller participant and one buyer participant are randomly paired. The pair is fixed **within the match**. After each match, participants will be reshuffled to form new pairs. You will not learn the identity of the participant you are paired with, nor will that participant learn your identity—even after the end of the experiment.

In this experiment, there is a seller who has an asset with no value, and a buyer who values the asset positively. The buyer's value of the asset is represented by $B$. At the start of each match, the computer randomly selects $B$ from the set $\{70, 240, 500\}$, where each value is equally likely to be chosen. $B$ is fixed for each match but is independently selected for each new match. Importantly, **the buyer knows the value $B$, but the seller does not.**

## Your Decisions in Each Match

Each match consists of up to **two rounds** of bargaining. In Round 1, a seller offers a price to sell the asset, and the buyer responds. If the offer is rejected, the match may move on to Round 2 of bargaining. The details follow.

**Your Task as a Seller in Round 1:** Suppose your role is a seller. At the beginning of Round 1, you will see the following figure. Three blue bars between 50 and 500 represent the possible values of $B$.

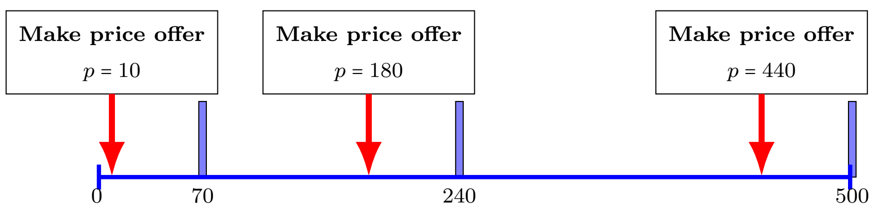


Figure 10: Seller's Screen: Round 1

Choose your price offer by clicking one of the three buttons: "Make price offer $p = 10$", "Make price offer $p = 180$", and "Make price offer $p = 440$". After that, click the submit button, and wait for the buyer's decision. You expect one of three possible outcomes.

- If the buyer **accepts** the offer, the match is over and you earn $p$ tokens.
- If the buyer **takes an outside option**, the match is over and you earn 0 tokens.
- If the buyer **rejects** the offer, then the match moves to Round 2 with an 80% chance. Note that if the match is terminated with a 20% chance, both you and the buyer earn 0 tokens.

When the buyer accepts the offer or takes an outside option, the buyer's decision will be correctly carried with a 99.9% chance. With a 0.1% chance (1 in 1,000), however, a server computer overrides the buyer's decision and rejects the offer. This minuscule probability is merely introduced to preserve a *theoretical* possibility of reaching Round 2. Since this probability is negligible, overriding is unlikely to happen in the course of your participation.

**Your Task as a Buyer in Round 1:** Suppose your role is a buyer. At the beginning of Round 1, you will see the following figure. The horizontal position of the blue bar represents $B$, your value of the asset. ($B$ is 240 in this example, but your value can be 70 or 500 as well.) Once the seller in your pair offers a price, $p$, a red vertical arrow will appear on the figure. The position of the red arrow represents $p$. After that, decide whether to

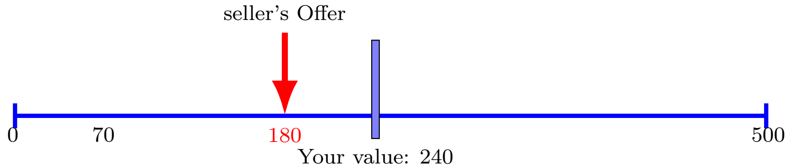


Figure 11: Buyer's Screen in Each Round

- **accept** the offer and earn $(B - p)$ (in which case the match is over),
- **reject** it and move on to the next round with an 80% chance, or
- **take an outside option** to earn 50 tokens (in which case the match is over).

Beware that you cannot accept the offered price $p$ if it is strictly greater than your value $B$, otherwise your payoff becomes negative. When you accept the offer or take an outside option, your decision will be correctly carried with a 99.9% chance. With a 0.1% chance (1 in 1,000), however, a server computer overrides your decision and rejects the offer. This minuscule probability is merely introduced to preserve a *theoretical* possibility of reaching Round 2. Since this probability is negligible, overriding is unlikely to happen in the course of your participation.

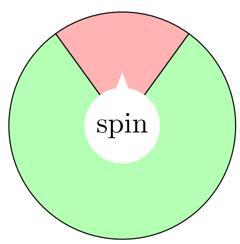


Figure 12: Spinning Wheel

**Transition to Round 2:** As described, when bargaining does not end in Round 1, the match continues to Round 2 with an 80% chance. Your screen presents a spinning wheel that consists of red area (20%) and green area (80%) as illustrated below. Once you click the Spin button, the wheel starts spinning. If the spinning wheel stops at the green area, the match continues. Otherwise, the match terminates. Note that the seller and buyer in the same match always see the same outcome from the wheel.

**Your Task as a Seller in Round 2:** Before submitting a new price offer, report how you believe the buyer's value, by filling out the following sentence.

| | | |
|---|---|---|
| | 70 with | a (____)% of chance, |
| I believe that the value of the buyer paired in this match is | 240 with | a (____)% of chance, |
| | 500 with | a (____)% of chance. |

The three numbers must sum up to 100. **The reported probabilities will appear in your decision screen but will not be shared with the buyer.**

Its sole objective is to help you think about an appropriate price offer. It is important to note that there are no advantages to indicating probabilities that differ from your true belief, so please report your belief as accurately as possible.

Note that the buyer's value of the asset ($B$) and the value of the outside option (50 tokens) will remain the same across rounds within a match.

**Your Task as a Buyer in Round 2:** Your screen will present the same figure as the one you had in Round 1 that indicates your value. Once the seller in your pair makes a price offer in Round 2, a red vertical arrow will appear on the figure. The position of the red arrow represents $p$. After that, decide whether to

- **accept** the offer and earn $(B - p)$,
- **reject** it and earn 0 tokens, or
- **take an outside option** to earn 50 tokens.

Beware that **the match will be over at the end of Round 2.**

## Information Feedback

- At the end of each **round**, you will know the seller's price offer and the buyer's decision. If the buyer rejects the offer, you will know whether the match is continued to the next round or terminated.
- At the end of each **match**, you will know how many tokens you receive from the match.

## Your Monetary Payments

At the end of the experiment, a computer will randomly select one match out of 10 for your payment. Every match has an equal chance to be selected for your payment, so it is in your best interest to take each match equally seriously. Participants will receive the amounts of tokens according to the outcome from the selected match with the exchange rate of 1 token = 1 HKD. Also, every participant will receive a **show-up fee of HKD 40**.

## Completion of the Experiment

After the 10th match, the experiment will be over. You will be instructed to fill in the receipt for your payment. The amount you earn will be paid **electronically via the HKUST Autopay System to the bank account you provide to the Student Information System (SIS)**. The Finance Office of HKUST will arrange the auto-payment. An email notification will be sent to your HKUST email address on the pay date under the name of the sender "FOPSAP" (Finance Office Payment System Auto Payment).

## Comprehension Check

To ensure your comprehension of the instructions, you will answer four multiple-choice questions. You can proceed only with all correct answers. Afterwards, you will participate in a practice match.

Q1 Suppose you are a seller. Which of the followings is NOT TRUE? (a) I do not know how much the buyer values the asset. (b) If the buyer takes an outside option, I earn 50 tokens. (c) If I offer 440 tokens, and the buyer accepts it, then I earn 440 tokens. (d) If bargaining does not end in Round 1, then I can make a new offer with a 80% chance.

Q2 Suppose you are a buyer, and the value of the asset is 500. Which of the followings is TRUE? (a) If I accept a price offer of 180 tokens, I earn 180 tokens. (b) If I take an outside option, I earn 550 tokens. (c) If I accept a price offer of 180 tokens, I earn 320 tokens. (d) In Round 2 of this match, the value of the asset will be different from 500.

Q3 Suppose the price offer in Round 1 is rejected. Which of the followings CAN HAPPEN? (a) The match is terminated, and both the seller and the buyer earn 0 tokens. (b) The match is continued forever, even after continuous rejections. (c) The match is terminated, and each participant in the pair earns a half of value $B$. (d) The match initiates an open chat to negotiate.

Q4 Suppose the first match is done. Which of the followings is TRUE? (a) It is almost sure that I will be paired with the same participant in the first match. (b) I may play another role different from what I did in the first match. (c) The buyer's value of the asset in the second match will be the same as the one in the first match. (d) My previous actions do not affect the value of the asset in the new match.

## 1 Practice Match and 10 Actual Matches

Thank you for paying attention to the instructions. Before you will play the 10 actual matches, you will have one practice match (Match #0) which is not relevant to your payment. Its objective is to get you familiar with the computer interface and the flow of the decisions in each round of a match. Once the practice match is over, it moves to the actual matches.