UTMD Working Paper
The University of Tokyo
Market Design Center

UTMD-116

The Upside of Unreliability: Screening via Sunk Costs in Platforms

Hiroaki Odahara
The University of Tokyo

February 12, 2026

The Upside of Unreliability: Screening via Sunk Costs in Platforms*

Hiroaki Odahara†

February 12, 2026

Abstract

Platforms often require users to incur sunk participation costs (fees, effort, or waiting) before learning whether service will be delivered. We study such environments with a model that separates actual non-delivery risk from perceived risk. When users are naïve about risk, unreliability creates scope for ex post manipulation: within a benchmark class of efficient two-bidder mechanisms, the war-of-attrition (second-price all-pay) format maximizes deviation revenue. When risk is common knowledge but budgets and valuations are misaligned, unreliability instead screens participation. Higher reliability can attract deep-pocket, lower-valuation entrants who crowd out budget-constrained high-valuation users, while intermediate unreliability can induce their exit and may increase consumer surplus discontinuously at the exit threshold. We do not advocate intentional unreliability; rather, we highlight a design tradeoff between trust-enhancing reliability and screening through endogenous entry in settings with unavoidable execution risk and limited commitment.

JEL Codes: D44, D47, D81, D82, L14
Keywords: platform reliability; non-delivery risk; sunk costs; all-pay auctions; exit-by-risk; budget constraints

1 Introduction

In the design of digital platforms, reliability is customarily regarded as a core requirement. From ride-sharing apps to decentralized finance (DeFi) protocols, operators strive to minimize friction and guarantee execution. However, many markets inherently require participants to incur *sunk costs* before an allocation is determined—whether it is the gas fees in a blockchain transaction, the waiting time for a ride that might be cancelled, the time spent in an online reservation queue that may crash or time out, or the non-refundable bidding fees in online auctions. In these settings, the risk of *non-delivery*—paying the cost but receiving nothing—is a fundamental feature of the market structure. While typically viewed as a source of inefficiency, this paper explores the counter-intuitive possibility that, when non-delivery risk is unavoidable and anticipated, its presence can sometimes serve a socially beneficial function by reshaping participation and competition in sunk-cost contests.

*This work was supported by JSPS KAKENHI Grant Number JP21H04979 and JST ERATO Grant Number JPMJER2301, Japan. Parts of this work were presented at the 2018 Spring Meeting of the Japan Economic Association (University of Hyogo). I thank Hitoshi Matsushima, Atsushi Iwasaki, Michihiro Kandori, Daisuke Oyama, Kenjiro Asami, Kazuyuki Higashi, Satoshi Kasamatsu, Daiki Kishishita, Satoshi Nakada, Munenori Nakasato, Shunya Noda, and Kyohei Okumura for helpful comments. Any remaining errors are my own.

†Graduate School of Economics, the University of Tokyo. hiroaki.odahara@e.u-tokyo.ac.jp

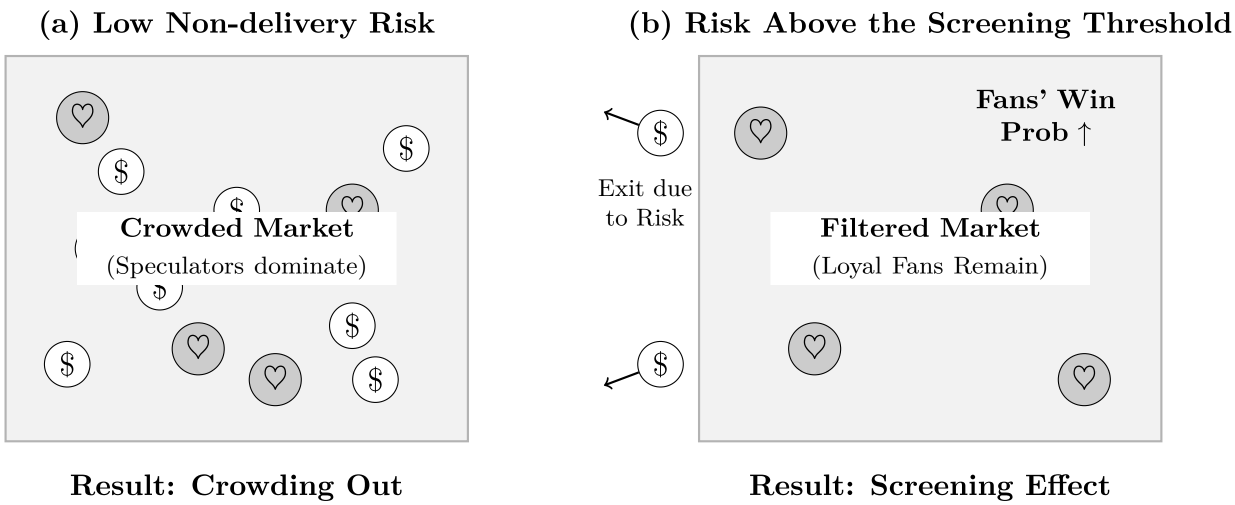


Figure 1: The Mechanics of Risk-Based Screening. In the reliable state (a), wealthy speculators ($) crowd out loyal fans (♡). In the unreliable state (b), speculators exit (or stay out) while loyal fans remain, thereby increasing the fans' allocation probability.

To understand how non-delivery risk can improve consumer surplus, consider membership programs that confer priority access to scarce future allocations (e.g., presale tickets, limited drops, or fan-club memberships). Participants can be heterogeneous: *high-valuation but budget-constrained* users ("loyal fans") coexist with *deep-pocket but lower-valuation* entrants ("speculators"/resellers). When non-delivery risk is low, speculators can profitably pay more (or wait longer) to dominate priority, crowding out budget-constrained high-valuation users. When there is a strictly positive chance that the sunk cost yields no allocation, aggressive spending becomes risky: if speculators must spend *more* to dominate, an intermediate level of non-delivery risk can make their expected return negative, inducing exit while leaving high-valuation users active. When delivery does occur, competition is less severe and the probability that high-valuation users obtain the allocation increases (Fig. 1). We refer to this mechanism as *exit-by-risk*: non-delivery risk acts as a screening force in sunk-cost competition.

This example also highlights a dual nature of non-delivery risk. On one hand, if participants are naïve about non-delivery risk, platforms can exploit sunk costs to extract excessive payments without delivering commensurate value—a mechanism reminiscent of low-trust "pay-to-bid" environments. On the other hand, if participants are rational and aware, the threat of non-delivery can correct market failures caused by budget constraints by deterring opportunistic deep-pocket entry. This raises a fundamental economic question: *under what conditions is a strictly positive probability of non-delivery ($q > 0$) socially beneficial, and when is it purely exploitative?*

Our key point is that the same sunk-cost primitive can produce opposite effects through a *belief wedge* ($\hat{q} - q$). When $\hat{q}$ understates true risk, non-delivery risk creates a channel for manipulation and revenue extraction. When $\hat{q} = q$, the same risk can become a strategic friction that reshapes participation, potentially improving outcomes under *budget–valuation mismatch*.

To analyze these forces within a single framework, we propose a unified model of sunk-cost competition: a war-of-attrition (second-price all-pay) auction with a reserve, augmented with a non-delivery state and belief-dependent behavior. This allows us to study two regimes within one

Table 1: Roadmap: trust deficits operate through manipulation under naïveté and screening under awareness.

	Naïve users ($\hat{q} = 0$) **ignore non-delivery**	**Aware users ($\hat{q} = q$)** **anticipate non-delivery**
Channel	Platform-side manipulation after bids are sunk	Exit-by-risk screens participation and competition
Outcome	Deviation revenue (exploitability), ERUS	Consumer surplus, CS
Headline result	WOA(0) attains the deviation-revenue upper bound $\mathbb{E}[v]$	CS may jump upward at the exit threshold $q^* = \underline{q}(v_H)$
Takeaway	High-powered sunk costs are most exploitable under naïveté	Reliability need not improve consumer surplus monotonically once entry is endogenous

reduced-form non-delivery model: a *naïve regime* ($\hat{q} = 0$) representing exploitative environments in which users ignore non-delivery risk, and an *aware regime* ($\hat{q} = q$) representing mature markets in which risk is common knowledge.

We correspondingly focus on two outcome concepts. In the naïve regime, we study the platform's potential for *deviation revenue*—revenue obtainable when a low-credibility platform can deny delivery or inject artificial competition *after* bids are sunk—and we summarize it using the benchmark objective ERUS (*expected revenue under unlimited seller-side deviation*), i.e., the expected payment extractable from a targeted bidder when the platform can force an effectively unbeatable rival. In the aware regime, we study *consumer surplus*, defined as the expected sum of bidders' utilities (including negative sunk payments), as a participant-side welfare proxy for how non-delivery risk affects allocation and rent dissipation; this differs from total surplus when sunk costs are transfers rather than deadweight losses. Because low-q equilibria can be non-unique, our aware-regime jump comparison is stated relative to a conservative dissipative benchmark class.

Table 1 provides a roadmap for the paper by contrasting how the same primitive—sunk-cost competition with a non-delivery state—operates through (i) a supply-side deviation channel in the naïve regime and (ii) a demand-side screening channel in the aware regime.

Here WOA(0) denotes our continuous penny-auction benchmark. The deviation-optimality result in Section 4 is specific to this benchmark (not to all-pay auctions), while Section 5 shows how common-knowledge risk can screen participation and generate an upward jump in CS at the exit threshold.

Naïve regime ($\hat{q} = 0$): exploitability and vulnerable formats. We first characterize which sunk-cost formats are most vulnerable to manipulation when users ignore non-delivery. Among a broad class of admissible efficient two-bidder mechanisms, we show that the WOA(0) format—our continuous penny-auction analogue—maximizes deviation revenue by attaining the ERUS upper bound (Theorem 4.2). We further show that deviation-optimal reserves are lower than standard revenue-optimal reserves (Proposition 4.6 and Example 4.7), mirroring the low posted starting prices commonly observed in fee-based competitive platforms.

The two-bidder benchmark is not only for tractability: it can be interpreted as a reduced-form representation of a platform-managed *two-player endgame*, in which users perceive a close race against a single rival. Keeping perceived rivalry small increases the marginal return to continuing to pay sunk costs, thereby creating especially favorable conditions for deviation revenue.

Aware regime ($\hat{q} = q$): screening via non-delivery risk and a consumer-surplus jump. Our main contribution concerns the aware regime and the possibility that non-delivery risk screens for intrinsic valuation. We study environments with budget–valuation mismatch, where high-valuation agents are liquidity constrained while deep-pocket agents have lower intrinsic valuations. In a canonical asymmetric instance, we identify a threshold non-delivery probability above which the deep-pocket low-valuation bidder exits in equilibrium ("exit-by-risk") while budget-constrained high-valuation bidders remain active (Proposition 5.2). This discontinuous change in participation can raise consumer surplus at the exit threshold, yielding an upward jump relative to the low-q dissipative benchmark (Proposition 5.10). We then provide a general screening result (order-statistic form) that isolates simple summary statistics of the effective budget distribution that govern both the exit cutoff and the surplus change (Proposition 5.14), showing that the mechanism is not an artifact of a knife-edge calibration. Figure 3b in Section 5 illustrates this piecewise behavior.

Our screening result relates to "ordeal mechanisms" that use costly frictions (e.g., waiting time) to induce self-selection. A distinctive feature here is that the friction is *probabilistic non-delivery risk* interacting with *sunk-cost competition*: because payments are sunk even when delivery fails, aggressive spending is penalized precisely on the margin where deep-pocket entrants must spend more to dominate. Importantly, our results do not advocate intentionally sabotaging delivery. Rather, they highlight that in environments where commitment is limited or execution cannot be perfectly guaranteed, reliability is not a monotone "more is always better" parameter once participation and budget constraints are endogenous.

These results offer a new perspective on platform design. While conventional wisdom suggests minimizing all friction, our findings imply that, when non-delivery risk is unavoidable, the interaction between sunk costs and that risk can act as a screening force that deters opportunistic arbitrage and improves assignment under budget constraints.

The remainder of the paper proceeds as follows. Section 2 discusses related literature. Section 3 presents the model (WOA(ε) with non-delivery and beliefs). Section 4 analyzes the naïve regime and characterizes formats and reserves that maximize deviation revenue. Section 5 studies the aware regime and establishes exit-by-risk screening and the resulting consumer-surplus jump, culminating in a general screening result. Section 6 discusses implications, limitations, and extensions.

2 Related Literature

Our work contributes to several strands of research on trust, frictions, and allocation in platform-mediated markets. A first point of contact is the growing literature on mechanisms that are *not* fully credible. Akbarpour and Li (2020) highlight that auctioneers often cannot commit to rules, creating a "trilemma" for mechanism design, and Komo et al. (2024) study shill-bidding as a profitable deviation strategy. We complement this line by focusing on *non-delivery* as a distinct deviation available to platforms. In particular, we characterize the optimal deviation revenue of an untrustworthy seller, which effectively captures a worst-case benchmark for user trust.

This credibility problem is especially salient in pay-to-bid and other all-pay-like environments, where participants often incur sunk costs before learning whether the platform will ultimately honor the announced allocation. Our "naïve regime" is motivated by evidence from penny auctions: Augenblick (2016) provides empirical evidence consistent with sunk-cost fallacy, while Hinnosaar (2016) offers a theoretical analysis and summarizes stylized institutional features such as low starting prices and heavy reliance on bid fees; Byers et al. (2010) further discuss information asymmetries in these environments. We formalize how such sunk costs can be exploited, but—crucially—contrast it with an "aware regime" in which rational users correctly perceive and respond to platform risk. Related work in computer science and AI also treats pay-to-bid platforms as data-rich algorithmic systems: bidder behavior can be predictable from large traces (Zhang et al., 2018), and learning-based models can forecast auction durations that govern bid-fee revenue (Wang and Yu, 2024).

Moving from behavioral motivations to screening incentives, our results in the aware regime relate to auctions and mechanism design with budget-constrained agents. Classic auction theory shows that budgets can overturn standard revenue and efficiency comparisons across formats (Che and Gale, 1998), and in algorithmic mechanism design, budgets motivate alternative welfare benchmarks; for example, Dobzinski and Paes Leme (2014) introduce *liquid welfare* and provide approximation guarantees. Our contribution is orthogonal: we show that non-delivery risk can act as an endogenous screening device when budgets and valuations are misaligned, generating discontinuous changes in participation and consumer surplus.

This screening interpretation also connects our analysis to the literature on welfare-improving frictions and "ordeal mechanisms" (Nichols and Zeckhauser, 1982). While ordeals typically impose deterministic costs (e.g., waiting time) to target the needy, our model highlights a different friction: *probabilistic* non-delivery risk. In competitive settings, such risk can selectively discourage wealthy speculators while preserving participation by budget-constrained high-valuation agents, showing how risk can succeed as a filter in environments where standard price mechanisms alone fail.

Finally, our motivating contrast between wealthy speculators and budget-constrained true fans links the paper to market design for tickets (e.g., Bhave and Budish, 2023) and to related "fan economy" environments modeled through all-pay-like competition (Tang et al., 2017). Unlike solutions that rely on price caps, lotteries, or identity verification, our mechanism leverages non-delivery risk and sunk costs to endogenously mitigate speculative entry, offering a complementary perspective on markets plagued by bots and scalpers. Technically, our analysis builds on the standard all-pay auction literature (Baye et al., 1996; Krishna and Morgan, 1997; Siegel, 2009) and on optimal auction theory more broadly (Myerson, 1981; Riley and Samuelson, 1981). Whereas these classic models assume perfect delivery, we extend the analysis to include objective and subjective non-delivery risks, capturing the "trust deficit" inherent in modern platforms.

3 Model

3.1 Agents and types

A seller offers a single indivisible good to a set of bidders $N = \{1, \ldots, n\}$. Bidder i has a private value $v_i \in [0, \bar{v}]$ for receiving the good and a private budget (or effort capacity) $c_i \in [0, \bar{c}] \cup \{\infty\}$ that limits how much she can ever pay. We write the type as $t_i = (v_i, c_i)$.

Unless stated otherwise, types are *independent across bidders* and *identically distributed* according to a common distribution H on $[0, \bar{v}] \times ([0, \bar{c}] \cup \{\infty\})$. The distribution H may allow arbitrary correlation between v_i and c_i (budgets need not be positively correlated with willingness to pay).[1]

[1] We keep an i.i.d. baseline for the main results, which yields clean equilibrium characterizations and aligns with standard revenue-equivalence arguments. In Section 5 we also study a stylized asymmetric instance (one deep-pocket/low-valuation bidder versus two budget-constrained/high-valuation bidders). This can be interpreted either literally (as bidders having different budget distributions across types, H_H and H_L) or as the conditional distribution

Utilities are quasi-linear. If bidder i receives the good and pays b_i, her utility is $u_i = v_i - b_i$; if she does not receive the good, $u_i = -b_i$. Bids are constrained by budgets: any feasible bid must satisfy $\beta_i \in [0, c_i]$.

3.2 Mechanism family: WOA with reserve

Fix a nonnegative reserve level $\varepsilon \geq 0$. A *war-of-attrition (second-price all-pay) auction with reserve*, hereafter referred to as WOA(ε), proceeds as follows. Each bidder simultaneously submits a bid $\beta_i \in [0, c_i]$.

- (**Eligibility**) A bidder is *active* if $\beta_i > \varepsilon$. Bidders with $\beta_i \leq \varepsilon$ are treated as not entering.
- (**Allocation**) If no bidder is active, no allocation occurs. Otherwise the good is allocated to an active bidder with the highest bid (ties broken uniformly at random among tied bidders).
- (**Payments**) If bidder i is inactive then $b_i = 0$. If bidder i is active and loses then she pays her bid, $b_i = \beta_i$. If bidder i is the winner then she pays the larger of the reserve and the highest losing bid:

$$b_i = \max\{\varepsilon,\ \max_{j \neq i} \beta_j\}.$$

Reserve boundary. We adopt a strict reserve: bids at the reserve are treated as non-entry, i.e., bidders with $\beta_i \leq \varepsilon$ are inactive. In particular, in WOA(0) bidding 0 is equivalent to abstaining (no chance of allocation).

For $\varepsilon = 0$, WOA(0) is a penny-auction analogue: a war-of-attrition (second-price all-pay) rule in bid space.

3.3 Non-delivery risk and beliefs

We augment WOA(ε) by a non-delivery state. Let $q \in [0, 1]$ be the *objective* probability that delivery fails or the platform overrides the advertised allocation after collecting participation costs.[2] Bidders may hold a belief $\hat{q} \in [0, 1]$ about this probability.

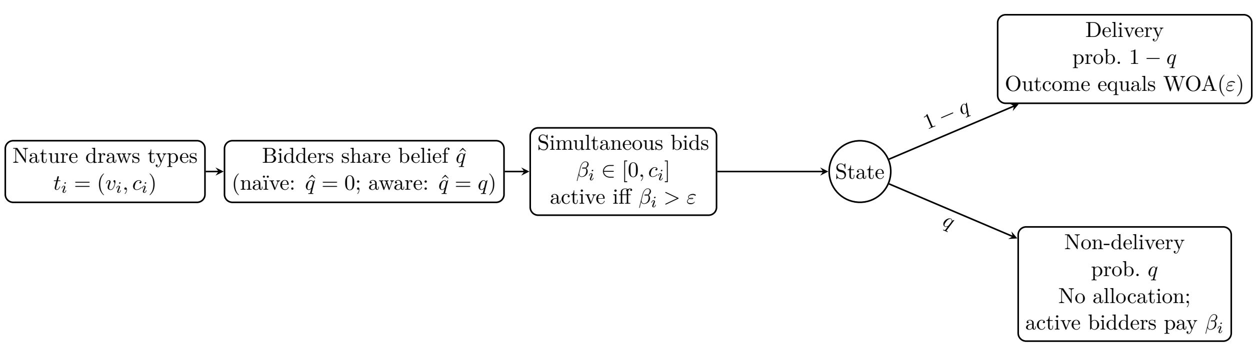


Figure 2: Timeline of WOA(ε) with non-delivery risk.

induced by an i.i.d. draw from a mixture distribution, given that the realized composition contains exactly one H-type and two L-types.

[2] In a queue interpretation, q is the probability that service is cancelled (or capacity is reduced) after agents have already sunk their waiting costs.

The timeline is illustrated in Figure 2: (i) bidders observe $\hat{q}$ and play a Bayes–Nash equilibrium of the perceived game, (ii) the non-delivery state realizes with probability q.

In the non-delivery state, no allocation occurs and all *active* bidders pay their bid: $b_i = \beta_i$ if $\beta_i > \varepsilon$, and $b_i = 0$ otherwise. In the delivery state (probability $1 - q$), the outcome coincides with WOA(ε).

Timing and interpretation. The delivery/non-delivery state is realized *after* bids are submitted but *before* the WOA(ε) allocation and (delivery-state) payment rule is executed. Thus bids can be interpreted as up-front sunk participation costs: when non-delivery occurs, the platform keeps these sunk costs and cancels without allocating or refunding. In particular, there is no winner in the non-delivery state, so no second-highest-bid payment is computed in that state. This timing is a stylized reduced form of queue/attrition environments, and it is the timing underlying the screening calculations in Section 5.

We focus on two information regimes:

- **Naïve regime:** $\hat{q} = 0$ (bidders ignore non-delivery risk).
- **Aware regime:** $\hat{q} = q$ (non-delivery risk is common knowledge).

3.4 Unlimited manipulation and the benchmark ERUS

In the naïve regime, a platform that can insert artificial competition and/or deny delivery can guarantee that a targeted bidder never wins, extracting from her the payment she would make when facing an "unbeatable" opponent. We summarize this manipulation revenue by a scalar objective.

Throughout Section 4 we restrict attention to $n = 2$ and ignore budgets (set $c_i = \infty$). Let $\bar{v}$ denote the top of the value support. Given a mechanism M with allocation/payment rules (y, b) and an equilibrium bidding function $\beta(\cdot)$ in the *perceived* game ($\hat{q} = 0$), define the *expected revenue under unlimited seller-side deviation* from bidder 1 as

$$\mathrm{ERUS}(M) \;\equiv\; \mathbb{E}_{v_1}\big[b_1(\beta(v_1),\, \beta(\bar{v}))\big], \tag{1}$$

where $\beta(\bar{v})$ represents the bid of an opponent with the highest possible valuation. Intuitively, this metric captures the expected payment collected from bidder 1 when the seller makes the opponent effectively "maximal".

4 Naïve Regime ($\hat{q} = 0$): Vulnerable Formats and Deviation Incentives

This section studies the *supply-side* channel of a trust deficit in sunk-cost competition. When participants behave as if delivery were guaranteed ($\hat{q} = 0$), a low-reliability platform that can deviate from the advertised outcome (after bids are sunk) has incentives to choose formats that collect payments even from losers. We summarize the resulting revenue opportunity by the benchmark objective ERUS(M) in (1) and refer to it as *deviation revenue*.

To keep the comparison sharp, throughout this section we impose three simplifying assumptions: (i) $n = 2$ bidders, (ii) no binding budgets (set $c_1 = c_2 = \infty$), and (iii) an *unrestricted* ability to deviate in the sense of (1). These assumptions are a benchmark rather than a claim about any particular market. In Appendix B we add a simple robustness extension showing that our mechanism ranking is unchanged when deviation is probabilistic or costly. Values are normalized to $v_i \in [0,1]$ with common CDF F and density f (continuous and positive on $(0,1)$).

Our baseline analyzes $n = 2$ to keep the equilibrium characterization and the reserve comparison transparent. This restriction is also economically natural in the manipulation scenarios we have in mind. In many sunk-cost platforms (including pay-to-bid environments), users do not condition their continuation decision on the full set of active participants; instead, they respond to perceived *effective rivalry* (e.g., "am I racing against one remaining bidder?").

When a platform can manipulate perceived competition (e.g., by injecting artificial activity), revenue extraction is often maximized by sustaining a *two-player endgame*. Keeping perceived rivalry small raises the perceived marginal return to continuing—each additional sunk payment has a higher chance of winning—thereby sustaining aggressive bidding and rent dissipation. We therefore view the $n = 2$ benchmark as a reduced-form representation of a platform-managed endgame and as a transparent worst-case benchmark for exploitability. A full analysis of the platform's optimal information policy over the number of remaining rivals is beyond the scope of this paper.

4.1 A benchmark mechanism class

We restrict attention to *full-allocation* efficient single-item auctions with a minimal "no-overcharging" payment discipline.

Definition 4.1 (Admissible efficient auctions)**.** A two-bidder mechanism $M = (y, b)$ is an *admissible efficient auction* if:

1. (**Full allocation & efficiency**) In any realized bid profile, the good is always allocated, and it is allocated to the bidder with the higher bid (ties broken arbitrarily).
2. (**No overcharging losers**) In any realized bid profile, the winner pays at least as much as the loser.
3. (**Normalization**) The lowest-value type makes zero expected payment in equilibrium.

Condition (1) fixes the allocation rule (no withholding upon delivery), so our format comparison holds allocation fixed. Condition (2) rules out mechanisms that penalize losing relative to winning, which could mechanically inflate ERUS but are atypical in implemented auction and platform formats. Condition (3) is the standard boundary condition needed for revenue-equivalence comparisons across efficient auctions. This definition is designed to isolate payment exposure (especially to losers) while holding allocation fixed, so that format comparisons speak directly to exploitability under sunk costs.

In what follows, when we evaluate ERUS(M) we focus on symmetric Bayes–Nash equilibria with strictly increasing bidding functions (the standard case under i.i.d. private values), so that the induced allocation rule is also efficient in values and the payment identity applies.

4.2 The war-of-attrition format WOA(0) maximizes deviation revenue

Having fixed the benchmark mechanism class, we can now ask which format is *most exploitable* when the platform can deviate after bids are sunk. The next theorem provides a sharp answer:

among admissible efficient auctions, the war-of-attrition (second-price all-pay) format WOA(0)—a penny-auction analogue with a zero reserve—maximizes deviation revenue ERUS.

Theorem 4.2 (War-of-attrition (second-price all-pay) maximizes ERUS in the full-allocation benchmark)**.** *Among admissible efficient auctions (Definition 4.1), evaluated at a symmetric equilibrium with a strictly increasing bidding function, the war-of-attrition (second-price all-pay) format* WOA(0) *maximizes* ERUS*. Moreover,*

$$\text{ERUS}(\text{WOA}(0)) = \mathbb{E}[v].$$

Intuition. Under any efficient two-bidder auction satisfying the normalization in Definition 4.1, revenue equivalence pins down the expected payment of the top type to be $\mathbb{E}[v]$. The "no overcharging losers" condition then implies that the payment extracted from a *targeted loser* under deviation cannot exceed $\mathbb{E}[v]$. The WOA(0) format attains this upper bound because it equalizes winner and loser payments pointwise, so extracting from the targeted loser is as profitable as extracting from the (effective) winner.

We first establish Lemma 4.3, which pins down the expected payment of the highest type via revenue equivalence. We then use it to bound ERUS for any admissible efficient auction and show that WOA(0) attains the bound, completing the proof of Theorem 4.2.

Lemma 4.3 (Payment identity at the top type)**.** *Consider any two-bidder mechanism M with independent private values $v_1, v_2 \sim F$ on $[0,1]$, risk-neutral bidders, and a symmetric equilibrium in which (i) bidding is strictly increasing in value and the object is always allocated to the higher bid, and (ii) the lowest type obtains zero expected utility (equivalently, makes zero expected payment). Then the expected payment of a bidder with value* 1 *equals* $\mathbb{E}[v]$.

Proof. See Appendix A.2. □

Sketch of proof. Fix any admissible efficient auction M (Definition 4.1) and a symmetric equilibrium with a strictly increasing bidding function.

Step 1 (revenue equivalence at the top type). By Lemma 4.3, the expected equilibrium payment of a bidder with value 1 equals $\mathbb{E}[v]$.

Step 2 (upper bound). Under unlimited manipulation, the seller can ensure bidder 1 never wins by making the opponent effectively maximal. Let $B_2(v_1)$ denote the (honest-play) equilibrium payment of the top type when the opponent has value v_1. The "no overcharging losers" condition in Definition 4.1(2) implies pointwise $b_1(\beta(v_1), \beta(1)) \leq B_2(v_1)$, and taking expectations yields $\text{ERUS}(M) \leq \mathbb{E}_{v_1}[B_2(v_1)] = \mathbb{E}[v]$.

Step 3 (tightness for WOA(0)*).* In WOA(0) with two bidders, winner and loser payments coincide in every realized profile (both equal the losing bid). Hence the bound in Step 2 binds, and $\text{ERUS}(\text{WOA}(0)) = \mathbb{E}[v]$.

See Appendix A.1 for the full proof. □

A robustness note: probabilistic or costly cancellation. Our benchmark objective ERUS(M) in (1) treats deviation (non-delivery) as unrestricted. Appendix B shows that the mechanism ranking in Theorem 4.2 is unchanged when deviation is limited. Specifically, suppose the platform can cancel with probability $\lambda \in [0,1]$ chosen ex ante and incurs a cost $k(\lambda)$, while bidders remain naïve about cancellation. Let $R(M)$ denote the platform's expected revenue in the delivery state under honest play of the advertised mechanism M. Then the platform's total expected payoff equals $(1-\lambda)R(M) + \lambda \cdot \mathrm{ERUS}(M) - k(\lambda)$. Since $R(M)$ is the same across admissible efficient auctions by revenue equivalence, the mechanism ranking for any fixed λ and cost function k is governed by the cancellation-state term. Accordingly, the platform's net payoff from cancellation states equals

$$\lambda \cdot \mathrm{ERUS}(M) - k(\lambda), \tag{2}$$

and maximizing the cancellation incentive over mechanisms is equivalent to maximizing $\mathrm{ERUS}(M)$. Hence WOA(0) remains optimal.

4.3 Deviation-optimal reserves are low

Theorem 4.2 compares formats within the full-allocation benchmark class in Definition 4.1. We now expand the design space by allowing the platform to impose a reserve $\varepsilon > 0$, which introduces a no-sale region (and hence falls outside the full-allocation restriction). This additional margin can increase ERUS—sometimes above $\mathbb{E}[v]$—at the cost of reduced allocative efficiency upon delivery, and therefore does not contradict Theorem 4.2.

To analyze reserve choices within the $\{\mathrm{WOA}(\varepsilon)\}$ family, we use the following equilibrium characterization.

Lemma 4.4 (Symmetric equilibrium in $\mathrm{WOA}(\varepsilon)$)**.** *Consider* $\mathrm{WOA}(\varepsilon)$ *with two bidders, values* $v \sim F$ *on* $[0,1]$ *with density* $f > 0$ *on* $(0,1)$*, and risk-neutral utilities. Define the cutoff* $v^* \in [0,1]$ *by*

$$v^* F(v^*) = \varepsilon. \tag{3}$$

There exists a symmetric equilibrium in which types $v \le v^*$ *bid* $\beta(v) = 0$*, and types* $v > v^*$ *bid*

$$\beta(v) = \varepsilon + \int_{v^*}^{v} \frac{t f(t)}{1 - F(t)}\, dt. \tag{4}$$

Proof. See Appendix A.3. □

Lemma 4.4 reduces the reserve-choice problem within the $\{\mathrm{WOA}(\varepsilon)\}$ family to the one-dimensional cutoff v^* (equivalently, ε), which we exploit below.

Let r_{ER} denote the standard revenue-maximizing reserve (equivalently, cutoff type) in a symmetric truthful efficient auction (Myerson–Riley–Samuelson). To compare reserves on the same scale as our $\mathrm{WOA}(\varepsilon)$ parameterization, we map this cutoff into the corresponding bid-threshold parameter

$$\varepsilon_{\mathrm{ER}} \equiv r_{\mathrm{ER}} F(r_{\mathrm{ER}}). \tag{5}$$

Let $\varepsilon_{\mathrm{ERUS}}$ maximize $\mathrm{ERUS}(\mathrm{WOA}(\varepsilon))$.

Definition 4.5 (Strictly monotone distributions)**.** A distribution F with density f on $[0,1]$ is *strictly monotone* if the function

$$g(v) \equiv 1 - F(v) - 2v f(v)$$

is strictly decreasing on $[0,1]$.

Proposition 4.6 (ERUS-optimal reserve is lower)**.** *If F is strictly monotone, then $\varepsilon_{\text{ERUS}} < \varepsilon_{\text{ER}}$.*

Intuitively, standard revenue maximization balances (i) extracting more from high types and (ii) losing trade when the reserve binds. Under the deviation objective ERUS, the platform monetizes *losing* payments (sunk costs) as well, which tilts the optimum toward admitting more types and hence toward a lower reserve.

Proof. See Appendix A.4. □

Example 4.7 (Uniform illustration)**.** *Suppose $v \sim \text{Unif}[0,1]$. Within the class $\{\text{WOA}(\varepsilon) : \varepsilon \in [0,1]\}$, the deviation-optimal cutoff is $v^* = 1/3$ and hence the deviation-optimal reserve is $\varepsilon^* = v^* F(v^*) = 1/9$. By contrast, the standard (reliability-aware) revenue-optimal cutoff under $v \sim \text{Unif}[0,1]$ is $r_{\text{ER}} = 1/2$. On the $\text{WOA}(\varepsilon)$ scale, this corresponds to $\varepsilon_{\text{ER}} = r_{\text{ER}} F(r_{\text{ER}}) = (1/2) \cdot (1/2) = 1/4$. Thus $\varepsilon^* = 1/9 < 1/4 = \varepsilon_{\text{ER}}$, consistent with Proposition 4.6.*

Documented implementations of pay-to-bid ("penny") auctions often feature very low posted starting prices, frequently zero, and derive an important share of revenue from non-refundable bid fees rather than the final winning price (Hinnosaar, 2016; Augenblick, 2016). Proposition 4.6 mirrors this structural feature: when the operator places weight on *deviation revenue*—revenue generated purely through sunk-cost payments in a non-delivery state—the optimal reserve is pushed toward its minimum. Importantly, low reserves are not diagnostic of deception: legitimate fee-based platforms may also choose zero reserves. Rather, our result explains why *low-reliability incentives* distort reserve choices downward relative to the standard revenue-optimal benchmark.

5 Awareness, Screening, and Consumer Surplus

Section 4 studied the naïve regime ($\hat{q} = 0$), where a trust deficit opens a supply-side deviation channel. In particular, within the full-allocation benchmark class, Theorem 4.2 shows that WOA(0) is an extreme point: it maximizes deviation revenue from sunk bid fees. We emphasize that this is a vulnerability diagnosis rather than an endorsement of penny auctions.

We now turn to the aware regime ($\hat{q} = q$) and hold the *sunk-cost competitive environment* fixed. When non-delivery occurs with probability q, the continuation value of winning is scaled by $(1-q)$, while sunk payments remain sunk. This wedge can change the set of active participants: in our asymmetric configuration (one deep-pocket/low-valuation bidder versus budget-constrained/high-valuation bidders), sufficiently high q induces an *exit-by-risk* of the deep-pocket bidder, which relaxes competitive pressure on the constrained high-valuation bidders. Because WOA(0) is a high-powered sunk-cost format, it can make this screening margin particularly salient: informally, the same property that makes it maximally exploitable under naïveté can make it a strong screening device under awareness.

We first characterize the exit-by-risk equilibrium and identify the threshold for the deep-pocket bidder to drop out (Proposition 5.2). We then show that this exit can generate a discontinuous jump in consumer surplus, relative to a dissipative benchmark in which the constrained bidders' surplus is fully dissipated when the deep-pocket bidder remains active (Proposition 5.10 and related results).

5.1 A canonical asymmetric instance with parameters

We work with a canonical asymmetric instance with one deep-pocket/low-valuation bidder and two budget-constrained/high-valuation bidders, mirroring the motivating queue story. There are three bidders $N = \{1, 2, 3\}$. Bidders 1 and 2 ("budget-constrained high-valuation") have a common valuation $v_L > 0$ for the good and private budgets $c_1, c_2 \sim \text{Unif}[0, 1]$ i.i.d.[3] Bidder 3 ("deep-pocket low-valuation") has a lower valuation $v_H \in (0, v_L)$ and an effectively unbounded budget ($c_3 = \infty$). All bidders observe q.

To keep the canonical instance focused on the budget-constrained screening mechanism, we concentrate on the region $v_L > 1$ and $v_H \geq 1$. The condition $v_L > 1$ makes the two-player subgame "budget-binding" in the sense of Lemma 5.1, so that budget-constrained bidders optimally bid their full budgets when bidder 3 exits. Meanwhile, $v_H \geq 1$ ensures that at low non-delivery risk bidder 3 has a strict incentive to bid 1 and guarantee winning against budget-constrained rivals (see (6) at $q = 0$). When either $v_L \leq 1$ or $v_H < 1$, equilibria can involve interior bidding and/or endogenous participation; see Appendix C for a general "bang–bang" reduced form in the L-only subgame.

The mechanism is WOA(0) with non-delivery probability q as in Section 3. Since $\varepsilon = 0$, every positive bid is active. If the non-delivery state occurs (probability q), no allocation occurs and each bidder pays her bid. Otherwise (probability $1 - q$), the object is allocated to the highest bidder; the winner pays the highest losing bid and each loser pays her own bid.

We interpret bids as *waiting* or *effort* costs: in both delivery and non-delivery states, a bidder who "stays" up to β_i pays β_i. The only difference is that in the non-delivery state nobody receives the good.

5.2 Equilibrium structure and exit-by-risk

We first note that if only bidders 1 and 2 participate, then bidding as much as the budget allows is optimal as long as q is not too high.

Lemma 5.1 (Two-player subgame). *Suppose bidder 3 bids $\beta_3 = 0$. If $q < \bar{q}(v_L) \equiv 1 - 1/v_L$, then $\beta_1 = c_1$ and $\beta_2 = c_2$ form a Bayes–Nash equilibrium for bidders 1 and 2.*

Proof. See Appendix A.5. □

Remark. The condition $q < \bar{q}(v_L) = 1 - 1/v_L$ is non-vacuous only when $v_L > 1$. If $v_L \leq 1$, then $\bar{q}(v_L) \leq 0$ and the lemma does not apply for any $q \in [0, 1]$.

We now ask whether bidder 3 wishes to enter. When the non-delivery probability is *intermediate*, bidder 3's expected benefit from outbidding budget-constrained rivals is outweighed by the expected rent dissipation (loss) from paying without delivery. In this region, an equilibrium features *exit* of the deep-pocket/low-valuation bidder.

Proposition 5.2 (Exit-by-risk equilibrium). *Define the cutoffs*

$$\underline{q}(v_H) \equiv 1 - \frac{1}{v_H + 1/3} \qquad \textit{and} \qquad \bar{q}(v_L) \equiv 1 - \frac{1}{v_L}.$$

[3] This normalization is without loss of generality for the comparative statics we highlight: a common scaling of all budgets and valuations rescales bids and payoffs proportionally.

If $q \in [\underline{q}(v_H), \bar{q}(v_L))$, then the profile

$$\beta_1 = c_1, \qquad \beta_2 = c_2, \qquad \beta_3 = 0$$

constitutes a Bayes–Nash equilibrium.

Intuition. When bidders 1 and 2 bid their budgets, bidder 3's best response is an endpoint: $\beta_3 \in \{0, 1\}$. His expected utility from bidding $\beta_3 = 1$ is

$$U_3(1) = (1-q)\Big(v_H + \tfrac{1}{3}\Big) - 1, \tag{6}$$

so he exits iff $U_3(1) \leq 0$, i.e., $q \geq \underline{q}(v_H)$. The constant $1/3$ reflects the order statistic of two budget draws and is not essential; see Propositions 5.5–5.6.

The interval $[\underline{q}(v_H), \bar{q}(v_L))$ is nonempty whenever $v_L > v_H + 1/3$. Economically, v_L must be sufficiently higher than v_H so that (i) budget-constrained bidders still prefer to "race" up to their budgets, while (ii) the deep-pocket bidder prefers to opt out once the dissipation risk q is large enough.

Sketch of proof. Fix $q \in [\underline{q}(v_H), \bar{q}(v_L))$.

Step 1. Given $\beta_3 = 0$, Lemma 5.1 implies that for $q < \bar{q}(v_L)$, bidders 1 and 2 best respond by $\beta_1 = c_1$ and $\beta_2 = c_2$.

Step 2. Given $\beta_1 = c_1$ and $\beta_2 = c_2$, bidder 3's expected payoff as a function of $\beta_3 \in [0, 1]$ is strictly convex, so his best response is an endpoint $\beta_3 \in \{0, 1\}$. Evaluating the endpoint payoff $U_3(1)$ yields (6), so bidder 3 exits iff $U_3(1) \leq 0 \iff q \geq \underline{q}(v_H)$.

Step 3. Combining Steps 1–2 gives the stated equilibrium for $q \in [\underline{q}(v_H), \bar{q}(v_L))$. See Appendix A.6 for the full proof. □

If instead non-delivery is very unlikely (q small), bidder 3 has a strong incentive to enter and outbid budget-constrained bidders. Then the pure-strategy equilibrium from Proposition 5.2 cannot exist.

Proposition 5.3 (No such pure equilibrium for low q)**.** *If $q \in (0, \underline{q}(v_H))$, there is no Bayes–Nash equilibrium in which $\beta_1(c_1) = c_1$ and $\beta_2(c_2) = c_2$ almost surely.*

Proof. See Appendix A.7. □

Remark 5.4 (Low-q equilibria)**.** For $q < \underline{q}(v_H)$, equilibria exist but may involve mixing and partial participation. Closed-form expressions depend on the distributional assumptions.

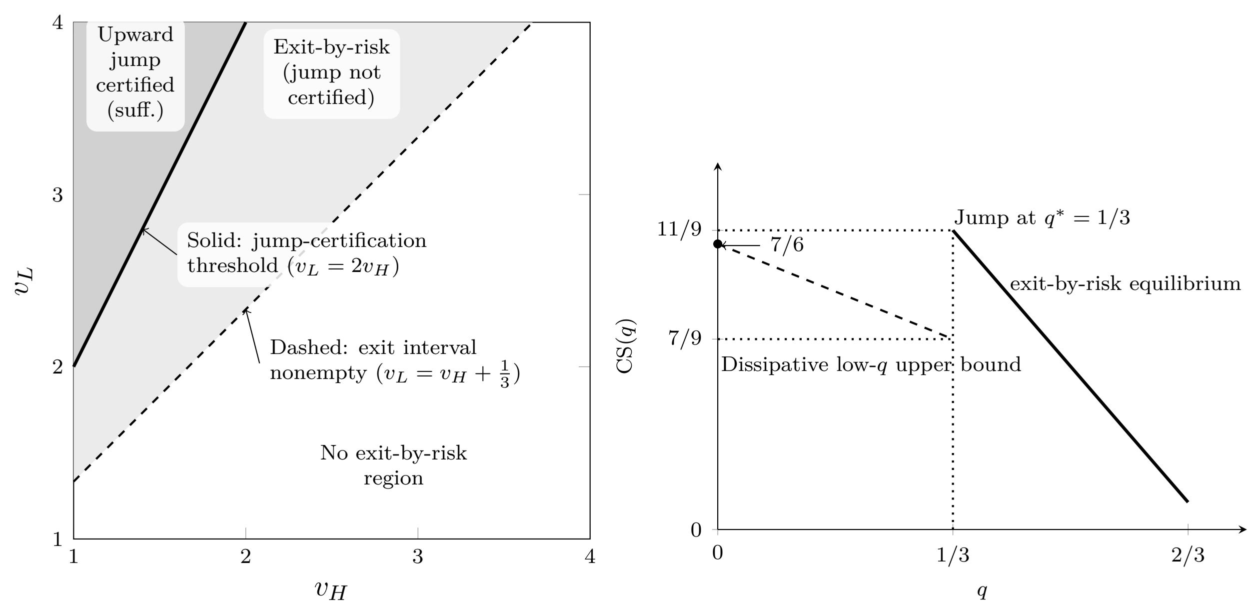
Accordingly, our consumer-surplus comparison below uses an equilibrium-class upper bound rather than a full characterization of low-q equilibria.

5.3 When does the consumer-surplus jump arise?

Having established when exit-by-risk occurs, we now ask when this participation shift translates into a consumer-surplus gain.

It is a generic consequence of two ingredients: *budget–valuation mismatch* and *differential sensitivity to wasted sunk costs.*

Figure 3a provides a simple regime map for the canonical instance, summarizing (i) when the exit-by-risk region is nonempty (Proposition 5.2) and (ii) when our bound-based comparison delivers a sufficient condition for an upward jump (Proposition 5.10) versus a sufficient condition under which the same bound comparison does not certify an *upward* jump (Corollary 5.12). To keep the picture focused on the budget-binding screening regime analyzed in this section, the map plots only $(v_H, v_L) \in [1,4] \times [1,4]$. Values below 1 are omitted because budget-constrained bids need not be budget-binding in that range, so the canonical screening comparison becomes sensitive to additional equilibrium cases.



(a) Canonical-instance regime map.

(b) Benchmark surplus jump.

Figure 3: Screening via non-delivery risk. Panel (a) plots a canonical-instance regime map over $(v_H, v_L) \in [1,4]^2$ (budget-binding window). The dashed boundary indicates when the exit-by-risk interval $[\underline{q}(v_H), \bar{q}(v_L))$ is nonempty. The solid boundary gives a *conservative sufficient condition* that certifies an upward consumer-surplus jump at $q^* = \underline{q}(v_H)$: dark shading marks the certified-jump region, while light shading marks parameter values where the exit-by-risk interval is nonempty but the bound comparison does not certify an upward jump (it may still occur). Panel (b) plots the consumer surplus in the exit-by-risk equilibrium (solid) together with the dissipative low-q upper bound from Lemma 5.9 (dashed) in the benchmark calibration $(v_L, v_H) = (3, 7/6)$. The vertical gap at $q^* = 1/3$ is the upward jump *certified by the bound comparison* in Proposition 5.10; consumer surplus then declines linearly over $[1/3, 2/3)$ within the exit-by-risk region. (see Example 5.13).

Panel (a) separates feasibility of exit-by-risk (dashed boundary) from our conservative certification of an upward jump (solid boundary). The jump is most likely when the type distribution features participants with *high budgets but low valuations* ("deep pockets") coexisting with participants with *high valuations but tight budgets*. Equivalently, budgets need not be positively correlated with willingness to pay. Such patterns can arise from negative valuation–income correlation, heterogeneity in liquidity constraints, or mixture populations with different tails.

Let q^* be the smallest non-delivery probability at which high-budget/low-valuation types exit while high-valuation budget-constrained types remain active. At q^*, the deep-pocket type is indifferent between entering and exiting, so her contribution to consumer surplus is (approximately) continuous. A positive jump can therefore be driven by an increase in the budget-constrained types' equilibrium utilities when the deep-pocket type exits. In the canonical instance, a sufficient condition is that consumer surplus *in the exit region* at q^* exceeds an upper bound on consumer surplus *below* q^*.

We make this logic explicit in the benchmark calibration in Proposition 5.10. Propositions 5.5 and 5.6 show that the key constants in the exit cutoff and in the exit-region consumer-surplus formula are not artifacts of a particular calibration: they generalize in simple ways to (i) more budget-constrained high-valuation bidders and (ii) other budget distributions.

Proposition 5.5 (More budget-constrained high-valuation bidders (K-generalization))**.** *Consider the variant with $K \geq 2$ budget-constrained high-valuation bidders with i.i.d. budgets* $\mathrm{Unif}[0,1]$ *who bid their budgets in the subgame where the deep-pocket bidder exits. Then for any bid $\beta \in [0,1]$, the deep-pocket bidder's expected utility against these rivals is*

$$U_H(\beta) = (1-q)\left(v_H \beta^K + \frac{1}{K+1}\beta^{K+1}\right) - \beta,$$

which is strictly convex on $[0,1]$. In particular, his expected utility from bidding 1 is

$$U_H(1) = (1-q)\Big(v_H + \tfrac{1}{K+1}\Big) - 1,$$

so his best response is an endpoint $\beta_H \in \{0,1\}$. Hence the exit threshold generalizes to

$$\underline{q}_K(v_H) = 1 - \frac{1}{v_H + 1/(K+1)}.$$

In particular, the constant $1/3$ in (6) corresponds to $K=2$ and arises from an order statistic rather than from a knife-edge calibration.

Proof. See Appendix A.8. □

Proposition 5.6 (Robustness to the budget distribution (moment form))**.** *Suppose bidders 1 and 2 have i.i.d. budgets $c_1, c_2 \sim G$ on $[0,1]$ and (for the relevant q range) bid $\beta_i = c_i$ whenever they participate. Then in the exit-by-risk region consumer surplus becomes*

$$\mathrm{CS}(q) = (1-q)v_L - \Big(q \cdot \mathbb{E}[c_1 + c_2] + (1-q)\cdot 2\,\mathbb{E}[\min\{c_1, c_2\}]\Big),$$

which is affine in q and depends on G only through the moments $\mathbb{E}[c_1 + c_2]$ and $\mathbb{E}[\min\{c_1, c_2\}]$. Hence the jump mechanism does not rely on uniform budgets; uniformity is used only to obtain *closed forms for the cutoffs and the illustration.*

Proof. See Appendix A.9. □

5.4 Consumer surplus and a jump at the exit threshold

Let $\mathrm{CS}(q)$ denote expected consumer surplus, defined as the expected sum of bidders' utilities (hence net of payments, and including negative sunk payments), as a function of q.

In the exit-by-risk region $q \in [\underline{q}(v_H), \bar{q}(v_L))$, only bidders 1 and 2 participate and bid their budgets. Conditional on a realized budget profile (c_1, c_2),

- in the non-delivery state, the total payment is $c_1 + c_2$;
- in the delivery state, the winner receives value v_L and the total payment is $2\min\{c_1, c_2\}$.

Taking expectations over $c_1, c_2 \sim \mathrm{Unif}[0,1]$ yields a simple affine expression:

$$\begin{aligned}\mathrm{CS}(q) &= (1-q)v_L \;-\; \Big(q \cdot \mathbb{E}[c_1 + c_2] + (1-q)\cdot 2\,\mathbb{E}[\min\{c_1, c_2\}]\Big) \\ &= (1-q)v_L \;-\; \Big(\tfrac{2}{3} + \tfrac{1}{3}q\Big), \qquad q \in [\underline{q}(v_H), \bar{q}(v_L)). \end{aligned} \tag{7}$$

In particular, within the exit-by-risk region, consumer surplus is strictly decreasing in q.

What matters for non-monotonicity is that the *identity of active bidders* can change discontinuously at $q = \underline{q}(v_H)$. As q crosses this threshold from below, the deep-pocket/low-valuation bidder exits, so the good is allocated (when delivered) to a high-valuation bidder, and the deep-pocket bidder's sunk payments disappear. This can generate an upward jump in consumer surplus.

Because equilibria need not be unique for $q < q^*$, we study whether this participation shift yields an *upward* discontinuity at $q^* = \underline{q}(v_H)$ using an equilibrium-class upper bound, rather than claiming a universal consumer-surplus ordering across all low-q equilibria.

Remark 5.7 (Where can dissipative low-q equilibria matter?)**.** In the uniform-budget canonical instance, the exit-by-risk equilibrium from Proposition 5.2 has bidder 3 inactive and bidders 1 and 2 playing the two-player subgame. For any $q < \bar{q}(v_L)$, a budget-constrained type with budget $c > 0$ obtains strictly positive expected utility when bidding $\beta = c$ against an opponent who bids her budget:

$$u_L(c) = (1-q)\Big(c(v_L - 1) + \tfrac{1}{2}c^2\Big) - qc = c\big((1-q)v_L - 1\big) + \tfrac{1-q}{2}c^2 > 0,$$

where the last inequality uses $q < \bar{q}(v_L) = 1 - 1/v_L$. Hence the exit-by-risk equilibrium is necessarily non-dissipative. The dissipative benchmark introduced below is therefore meant to capture an extreme low-risk regime $q < q^*$ in which bidder 3 remains active and competitive pressure can in principle drive the constrained bidders' surplus close to zero. We do not attempt to characterize equilibrium selection for each $q < q^*$.

To obtain a clean low-q benchmark without solving the full equilibrium correspondence, we introduce a conservative dissipative class.

Definition 5.8 (Dissipative low-q equilibria)**.** Fix $q < q^*$. We call a Bayes–Nash equilibrium *dissipative* if bidders 1 and 2 obtain zero expected utility. This is a conservative benchmark capturing the extreme case of full rent dissipation by the constrained bidders when bidder 3 remains active.

Lemma 5.9 (Low-q upper bound in dissipative equilibria)**.** *Fix any parameters* (v_L, v_H) *and any* $q < q^*$. *In any dissipative equilibrium (Definition 5.8), consumer surplus satisfies*

$$\mathrm{CS}(q) \le (1-q)v_H.$$

Intuitively, dissipative means the constrained bidders contribute no utility, so consumer surplus can be bounded by the deep-pocket bidder's delivery value, at most $(1-q)v_H$.

For convenience, define the *dissipative benchmark upper bound* by

$$\overline{\mathrm{CS}}^{\mathrm{diss}}(q) := (1-q)v_H.$$

Lemma 5.9 shows that $\mathrm{CS}(q) \le \overline{\mathrm{CS}}^{\mathrm{diss}}(q)$ in any dissipative equilibrium.

Proof. See Appendix A.10. □

Proposition 5.10 (A sufficient condition for an upward consumer-surplus jump)**.** *Suppose* $v_H > 2/3$ *so that* $q^* = \underline{q}(v_H) \in (0,1)$. *In the canonical instance with* $c_1, c_2 \sim \mathrm{Unif}[0,1]$, *if*

$$v_L > 2v_H, \tag{8}$$

then the exit-by-risk equilibrium at $q = q^*$ *delivers consumer surplus strictly above the low-*q *upper bound from Lemma 5.9. More precisely, in the exit-by-risk equilibrium at* q^*,

$$\mathrm{CS}(q^*) - (1-q^*)v_H \;\ge\; \frac{v_L - 2v_H}{v_H + 1/3} \;>\; 0,$$

and therefore $\mathrm{CS}(q^*) > \overline{\mathrm{CS}}^{\mathrm{diss}}(q^*)$. *In particular, for any sequence of dissipative equilibria for* $q < q^*$, *we have* $\mathrm{CS}(q^*) > \limsup_{q\uparrow q^*} \mathrm{CS}(q)$.

Proof. See Appendix A.11. □

Remark 5.11 (A slackened low-q upper bound)**.** Let $\mathrm{CS}(q)$ denote consumer surplus in any equilibrium at $q < q^*$, and define the aggregate utility of the constrained bidders by $R(q) \equiv U_1(q) + U_2(q)$. Since bidder 3 obtains value only upon delivery and payments are nonnegative in every state, we always have $U_3(q) \le (1-q)v_H$. Hence, in any equilibrium,

$$\mathrm{CS}(q) = R(q) + U_3(q) \;\le\; (1-q)v_H + R(q).$$

Consequently, Proposition 5.10 implies an upward jump at q^* for any sequence of low-q equilibria satisfying

$$\limsup_{q\uparrow q^*} R(q) \;<\; \frac{v_L - 2v_H}{v_H + 1/3}.$$

Corollary 5.12 (A sufficient condition for *no* upward jump in our comparison)**.** *In the canonical instance with* $c_1, c_2 \sim \mathrm{Unif}[0,1]$, *if* $v_L \le 2v_H$ *then*

$$\mathrm{CS}(q^*) \le (1-q^*)v_H \qquad \textit{in the exit-by-risk equilibrium at } q^* = \underline{q}(v_H).$$

Consequently, the exit-by-risk equilibrium at q^* cannot yield consumer surplus strictly above the low-q upper bound from Lemma 5.9; in particular, the bound-based argument for an upward *jump fails.*

Proof. See Appendix A.12. □

Example 5.13 (Benchmark calibration)**.** *Take $(v_L, v_H) = (3, 7/6)$. Then $q^* = \underline{q}(v_H) = 1/3$ and the exit-by-risk region is $[1/3, 2/3)$. In that region, (7) yields $\mathrm{CS}(q) = \frac{7}{3} - \frac{10}{3}q$, so $\mathrm{CS}(1/3) = 11/9$. Moreover, $v_L > 2v_H$ holds, and the bound from Lemma 5.9 implies $\limsup_{q\uparrow 1/3} \mathrm{CS}(q) \le (1-1/3)v_H = 7/9$ for dissipative low-q equilibria. Thus consumer surplus jumps upward by at least $4/9$ at $q = 1/3$ relative to the dissipative low-q benchmark.*

5.5 A general screening result (order-statistic form)

The preceding results can be summarized in a compact "general result" that isolates the few statistics that drive both the exit cutoff and the consumer-surplus jump. We emphasize that this result is a *conditional summary*: it does not attempt to characterize equilibria for arbitrary budget primitives, but shows that whenever an exit-by-risk region with budget-bidding exists, the relevant payoff objects depend on the (induced) budget distribution only through simple *order-statistic moments*.

Budget constraints can affect not only bids but also *participation*: some high-valuation bidders may optimally abstain and bid 0. Accordingly, in Proposition 5.14 we interpret G as the distribution of *effective bids* in the L-subgame, allowing an atom at 0 as a reduced-form representation of non-participation.

Appendix C records a general payoff expression for the L-only subgame and gives a simple sufficient condition under which best responses are *bang–bang* (bid either 0 or the full budget). Under our strict reserve convention (Section 3), bidding 0 is true non-entry and yields zero payoff. For active positive-budget L-types in continuous bid spaces, the symmetric L-subgame therefore places no atom at 0 when the all-zero event has positive probability (Remark C.2). In particular, for $K = 2$ and $q < \bar{q}(v_L)$ (Lemma 5.1), all L-types bid their budgets.

Proposition 5.14 (Screening via non-delivery risk: an order-statistic summary)**.** *Consider* WOA(0) *in the aware regime $\hat{q} = q$. Suppose there are $K \ge 2$ "high-valuation" bidders (the L-group) with common valuation v_L and i.i.d. budgets $c_1, \dots, c_K \sim G$ supported on $[0, 1]$ (allowing an atom at 0), and one "deep-pocket" bidder H with value v_H and an unbounded budget. Assume there is a q-interval I on which (i) when H bids 0, the L-bidders form a Bayes–Nash equilibrium by bidding their budgets, $\beta_i = c_i$, and (ii) against budget-bidding rivals, bidder H's best response is an endpoint $\beta_H \in \{0, 1\}$.*

Let $C_{(1)} \le \cdots \le C_{(K)}$ denote the order statistics of $(c_1, \dots, c_K)$ and define the three summary statistics

$$\mu_{\mathrm{sum}} \equiv \mathbb{E}\Big[\sum_{i=1}^{K} c_i\Big], \qquad \mu_{\max} \equiv \mathbb{E}[C_{(K)}], \qquad \mu_{\mathrm{gap}} \equiv \mathbb{E}\big[C_{(K)} - C_{(K-1)}\big].$$

Here $\mu_{\max}$ captures the typical strongest budget the deep-pocket bidder must beat (and hence the expected payment upon delivery), while μ_{gap} captures the expected winning margin that governs second-price/all-pay dissipation. The term μ_{sum} summarizes the expected sunk payments in the non-delivery state.

(i) ***(Exit cutoff).*** *Bidder* H*'s expected utility from bidding* 1 *against budget-bidding* L*-bidders is*

$$U_H(1) = (1-q)(v_H - \mu_{\max}) - q.$$

If $v_H > \mu_{\max}$ *and* $q \geq q^*$ *where*

$$q^* \equiv \frac{v_H - \mu_{\max}}{1 + v_H - \mu_{\max}},$$

then $U_H(1) \leq 0$*, and under the endpoint-best-response property* H *exits (bids* 0*). Hence an exit-by-risk equilibrium exists for all* $q \in [q^*, I)$*.*

(ii) ***(Consumer surplus in the exit equilibrium).*** *In the exit-by-risk equilibrium, expected consumer surplus is affine in* q *and admits the order-statistic form*

$$\mathrm{CS}_{\mathrm{exit}}(q) = (1-q)v_L - \mu_{\mathrm{sum}} + (1-q)\mu_{\mathrm{gap}}.$$

(iii) ***(A jump condition relative to dissipative low-*** q ***equilibria).*** *Fix* $q < q^*$ *and consider any dissipative equilibrium in the sense of Definition 5.8. Then* $\mathrm{CS}(q) \leq (1-q)v_H$*. Consequently, consumer surplus jumps upward at* $q = q^*$ *(relative to the dissipative upper bound) whenever*

$$\mathrm{CS}_{\mathrm{exit}}(q^*) > (1-q^*)v_H \qquad \iff \qquad v_L - v_H + \mu_{\mathrm{gap}} > \mu_{\mathrm{sum}}(1 + v_H - \mu_{\max}).$$

If the inequality fails (weakly), then $\mathrm{CS}_{\mathrm{exit}}(q^*)$ *does not exceed the dissipative bound, so the bound-based argument for an* upward *jump cannot be invoked.*

Specializing to $K = 2$ and $G = \mathrm{Unif}[0,1]$ recovers $\mu_{\mathrm{sum}} = 1$, $\mu_{\max} = 2/3$, and $\mu_{\mathrm{gap}} = 1/3$, matching the canonical cutoffs and the linear surplus formula above.
Proof. See Appendix A.13.

6 Discussion

6.1 Two Channels of Trust Deficits

This paper highlights a tension in markets where participants incur sunk costs before allocation and delivery is imperfect. In the naïve regime ($\hat{q} = 0$), a trust deficit creates a *supply-side deviation channel*: formats that collect sunk fees from many participants become particularly vulnerable when users underestimate non-delivery risk. In the aware regime ($\hat{q} = q$), the same primitive can generate a *demand-side screening channel*: when budgets and valuations are misaligned, non-delivery risk changes who enters and can improve assignment. Importantly, our analysis is *not* a prescription to introduce arbitrary non-delivery risk. Instead, it clarifies (i) when non-delivery risk is harmful because it opens a deviation channel, and (ii) when "effective frictions" can improve assignment by inducing self-selection among heterogeneous participants.

6.2 Design and Policy Implications

Section 4 suggests a diagnostic for platform fragility. When fees are largely sunk and reserve prices are pushed down, perceived competition can be sustained even when the platform cannot credibly commit to delivery. This points to consumer-protection levers that target the deviation channel, especially when users systematically underweight non-delivery risk. Examples include:

- **Verifiable reliability disclosure**: publish and audit non-delivery/cancellation rates and complaint-resolution statistics.
- **Commitment devices**: escrow, chargeback-compatible payment rails, or refund rules contingent on delivery.
- **Design choices that reduce post-bid deviation**: limiting discretionary cancellation and clarifying verification criteria ex ante.

Section 5 offers a complementary lesson for legitimate allocation problems with budget constraints. When wealthy but low-valuation agents crowd out budget-constrained high-valuation agents, screening can be valuable. Crucially, the screening logic does *not* require opaque non-delivery: the same self-selection force can be implemented with *transparent* and enforceable frictions, so that "screening" need not come at the expense of trust.

The right choice depends on what correlates least with budgets and most with the designer's notion of "commitment" or "intrinsic valuation." Examples include:

- **Refundable deposits / liquidity locks**: participants post a deposit that is returned upon delivery; the opportunity cost of locked funds can deter low-valuation arbitrage.
- **Non-transferability and anti-resale design**: identity checks at redemption, non-transferable access tokens, or mechanisms that reduce secondary-market profits.
- **Ordeal mechanisms that are publicly specified**: waiting times or queues with transparent rules (rather than hidden non-delivery risk), possibly combined with verification.
- **Proof-of-personhood / anti-bot frictions**: rate limits, verification steps, or KYC-type checks that disproportionately burden automated/speculative entry.

The key design question is which friction screens opportunistic participation *without* deterring high-valuation users who may be budget constrained, and how the platform can implement that friction while maintaining credible delivery commitments.

6.3 Empirical Predictions

Our theory yields several testable implications. First, in environments with weak trust and user naïveté, fee-driven formats should feature very low posted prices and revenue concentrated in sunk fees, alongside heightened sensitivity to perceived rivalry (e.g., platform UI features that emphasize a small number of "remaining" rivals). Second, in settings with budget–valuation mismatch, changes in perceived non-delivery risk should induce *composition shifts* in participation: a sufficiently high non-delivery probability (or a transparent substitute friction) can trigger exit by deep-pocket, low-valuation agents, leading to a discrete increase in the winning prospects of high-valuation agents conditional on delivery. Third, the two-channel perspective predicts a *potentially non-monotone* relationship between measured non-delivery risk and consumer outcomes across market maturity: reducing non-delivery risk is unambiguously beneficial when it closes the belief wedge, but can remove a screening force when the main distortion is budget-driven misallocation.

6.4 Limitations and Extensions

Our aware-regime analysis relies on a stylized form of budget–valuation mismatch, and the exit-by-risk mechanism may not arise when budgets and valuations are positively correlated. On the strategic side, we take beliefs as exogenous ($\hat{q} \in \{0, q\}$) and model the two-player endgame in Section 4 in reduced form. Natural extensions include learning about non-delivery risk, costly or bounded manipulation, richer bidder populations, and welfare benchmarks that incorporate platform surplus.

A further limitation is equilibrium multiplicity outside the exit-by-risk region. For low non-delivery risk (small q), entry incentives can sustain multiple participation patterns (including mixing or partial participation), and we do not characterize the full set of low-q equilibria or provide a general selection theory. Accordingly, our "jump" and non-monotone comparative-static claims in non-delivery risk are stated relative to a conservative benchmark class of low-q outcomes that isolates rent dissipation and the screening force; different equilibrium selection in practice could change the magnitude—and potentially the sign—of the welfare comparisons. Developing equilibrium-selection foundations (e.g., via learning, dynamics, perturbations, or institutional rules that pin down participation) is an important direction for future work.

Our welfare comparisons focus on consumer surplus (the sum of bidders' utilities), treating sunk payments as losses to participants. In applications, these sunk costs may be transfers (e.g., membership fees) or real resource costs (e.g., time spent waiting), so mapping our results to total surplus requires application-specific accounting. A unified treatment of transfers versus deadweight frictions—and the resulting design implications for non-delivery risk and screening—is another important direction for future work.

6.5 Broader Applicability

Beyond auctions, the insight that "risk screens for valuation" applies to decentralized systems and queueing mechanisms. In blockchain protocols, for instance, the cost associated with potential transaction failure functions as a sunk cost that prioritizes urgent transactions. In public service allocation, "ordeal" mechanisms such as waiting times can improve targeting by inducing self-selection among recipients. More broadly, our framework provides a lens to analyze how non-delivery risk, explicit screening frictions, and equilibrium selection jointly shape allocative efficiency in modern digital ecosystems.

References

Mohammad Akbarpour and Shengwu Li. Credible auctions: A trilemma. *Econometrica*, 88(2): 425–467, 2020. doi: 10.3982/ECTA15925.

Ned Augenblick. The sunk-cost fallacy in penny auctions. *The Review of Economic Studies*, 83(1): 58–86, 2016. doi: 10.1093/restud/rdv037.

Michael R. Baye, Dan Kovenock, and Casper G. de Vries. The all-pay auction with complete information. *Economic Theory*, 8(2):291–305, 1996. doi: 10.1007/BF01211819.

Aditya Bhave and Eric Budish. Primary-market auctions for event tickets: Eliminating the rents of "bob the broker"? *American Economic Journal: Microeconomics*, 15(1):142–170, 2023. doi: 10.1257/mic.20180230.

John W. Byers, Michael Mitzenmacher, and Georgios Zervas. Information asymmetries in pay-per-bid auctions. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC '10)*, pages 1–12. ACM, 2010. doi: 10.1145/1807342.1807344.

Yeon-Koo Che and Ian Gale. Standard auctions with financially constrained bidders. *The Review of Economic Studies*, 65(1):1–21, 1998. doi: 10.1111/1467-937X.00033.

Shahar Dobzinski and Renato Paes Leme. Efficiency guarantees in auctions with budgets. In *Automata, Languages, and Programming (ICALP 2014)*, volume 8572 of *Lecture Notes in Computer Science*, pages 392–404. Springer, 2014. doi: 10.1007/978-3-662-43948-7_33.

Toomas Hinnosaar. Penny auctions. *International Journal of Industrial Organization*, 48:59–87, 2016. doi: 10.1016/j.ijindorg.2016.06.005.

Andrew Komo, Scott Duke Kominers, and Tim Roughgarden. Shill-proof auctions. arXiv preprint, 2024.

Vijay Krishna and John Morgan. An analysis of the war of attrition and the all-pay auction. *Journal of Economic Theory*, 72(2):343–362, 1997. doi: 10.1006/jeth.1996.2208.

Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981. doi: 10.1287/moor.6.1.58.

Albert L Nichols and Richard J Zeckhauser. Targeting transfers through restrictions on recipients. *The American Economic Review*, 72(2):372–377, 1982.

John G. Riley and William F. Samuelson. Optimal auctions. *American Economic Review*, 71(3): 381–392, 1981.

Ron Siegel. All-pay contests. *Econometrica*, 77(1):71–92, 2009. doi: 10.3982/ECTA7537.

Pingzhong Tang, Yulong Zeng, and Song Zuo. Fans economy and all-pay auctions with proportional allocations. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI '17)*, pages 713–719. AAAI Press, 2017.

Yujia Wang and Haoran Yu. Predicting real-world penny auction durations by integrating game theory and machine learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38 (9):9926–9934, 2024. doi: 10.1609/aaai.v38i9.28854.

Xinyi Zhang, Shawn Shan, Shiliang Tang, Haitao Zheng, and Ben Y. Zhao. Penny auctions are predictable: Predicting and profiling user behavior on dealdash. In *Proceedings of the 29th ACM Conference on Hypertext and Social Media (HT '18)*, pages 123–127. ACM, 2018. doi: 10.1145/3209542.3209576.

A Proofs omitted from the main text

A.1 Proof of Theorem 4.2

Proof. Let M be any admissible efficient auction (Definition 4.1) with equilibrium bidding function $\beta(\cdot)$.

Step 1: an upper bound from revenue equivalence. Fix bidder 2's value at the top type $v_2 = 1$. Under the symmetric equilibrium with a strictly increasing bidding function, the top type bids highest, so bidder 2 wins the good for every realization of v_1 (ties have probability zero). Let $B_2(v_1)$ denote bidder 2's equilibrium payment in this event.

By Lemma 4.3,

$$\mathbb{E}_{v_1}[B_2(v_1)] = \mathbb{E}[v]. \tag{9}$$

Step 2: bound deviation revenue by the highest type's payment. Under unlimited manipulation, the seller can ensure bidder 1 never wins by making the opponent effectively "maximal." By definition,

$$\text{ERUS}(M) = \mathbb{E}_{v_1}[b_1(\beta(v_1), \beta(1))].$$

In every realized bid profile in which bidder 2 is maximal and thus wins, the "no overcharging losers" condition in Definition 4.1(2) implies

$$b_1(\beta(v_1), \beta(1)) \ \leq \ b_2(\beta(v_1), \beta(1)) \equiv B_2(v_1).$$

Taking expectations over v_1 and applying (9) yields

$$\text{ERUS}(M) \leq \mathbb{E}_{v_1}[B_2(v_1)] = \mathbb{E}[v].$$

Step 3: the bound is tight for the war-of-attrition (penny-auction analogue) format. Under WOA(0) with two bidders, the winner pays the losing bid and the loser pays her own bid. Hence in every realized profile, the winner and loser make the same payment (equal to the losing bid). Therefore, when bidder 2 is maximal, bidder 1's payment equals bidder 2's payment pointwise:

$$b_1^{\text{WOA}(0)}(\beta(v_1), \beta(1)) = b_2^{\text{WOA}(0)}(\beta(v_1), \beta(1)).$$

Consequently,

$$\text{ERUS}(\text{WOA}(0)) = \mathbb{E}_{v_1}\left[b_2^{\text{WOA}(0)}(\beta(v_1), \beta(1))\right].$$

Applying Lemma 4.3 to WOA(0) (which is efficient and satisfies the boundary condition) gives $\text{ERUS}(\text{WOA}(0)) = \mathbb{E}[v]$, and the upper bound is attained.

Combining Steps 1–3 shows that WOA(0) maximizes ERUS among admissible efficient auctions and achieves $\mathbb{E}[v]$. □

A.2 Proof of Lemma 4.3

Proof. Under the assumption that the equilibrium bidding function is strictly increasing (and continuity of F), bids are strictly ordered by values, so a bidder with value v wins if and only if the opponent's value is below v (ties occur with probability zero). Hence the allocation probability is $x(v) = F(v)$. The standard envelope/payment-identity argument implies $U'(v) = x(v) = F(v)$ and, with $U(0) = 0$, $U(v) = \int_0^v F(t)\,dt$. Expected payment is therefore $p(v) = vx(v) - U(v) = vF(v) - \int_0^v F(t)\,dt$. Evaluating at $v = 1$ gives $p(1) = 1 - \int_0^1 F(t)\,dt = \int_0^1 (1 - F(t))\,dt = \mathbb{E}[v]$. □

A.3 Proof of Lemma 4.4

Proof. We outline the standard derivation (war-of-attrition / second-price all-pay with a reserve).

Fix an increasing opponent strategy $\beta(\cdot)$ and consider a type v who submits a bid $b > \varepsilon$. Let V denote the opponent's value and $B = \beta(V)$ her bid. Conditional on $B < b$, bidder 1 wins, receives value v, and pays $\max\{\varepsilon, B\}$ (the larger of the reserve and the highest losing bid). Conditional on $B > b$, bidder 1 loses and pays her own bid b.

Thus expected utility from bid $b > \varepsilon$ is

$$U(v, b) = \mathbb{P}(B < b) \cdot v \;-\; \mathbb{E}[\max\{\varepsilon, B\} \cdot \mathbf{1}\{B < b\}] \;-\; b \cdot \mathbb{P}(B > b).$$

(In ties we can assign either bidder as winner; this does not affect the argument under continuous types.)

A note on the reserve term. For bids $b > \varepsilon$, the difference between the expression above and the "WOA(0)-style" formula that replaces $\max\{\varepsilon, B\}$ by B is a constant in v and b (it depends only on the opponent's bid distribution). Therefore the envelope condition and the differential equation characterization below are unchanged.

Let G be the CDF of bids induced by $\beta(\cdot)$. Since β is increasing, $G(b) = F(\beta^{-1}(b))$ and $B < b$ is equivalent to $V < \beta^{-1}(b)$.

A standard envelope argument (equivalently, indifference of types along the support under monotone strategies) implies that in a symmetric equilibrium, the equilibrium utility $U(v) \equiv U(v, \beta(v))$ satisfies

$$U'(v) = \mathbb{P}(V < v) = F(v), \tag{10}$$

with boundary condition $U(v^*) = 0$ at the cutoff type who is indifferent between entering ($b = \varepsilon^+$) and not entering ($b = 0$). Solving (10) gives

$$U(v) = \int_{v^*}^{v} F(t)\,dt \qquad \text{for } v \geq v^*.$$

On the other hand, substituting $b = \beta(v)$ into the payoff formula and using symmetry yields

$$U(v) = F(v) \cdot v \;-\; \int_0^v \beta(t)\,dF(t) \;-\; \beta(v) \cdot (1 - F(v)) \;-\; C, \qquad C \equiv \mathbb{E}[(\varepsilon - B) \cdot \mathbf{1}\{B < \varepsilon\}]\,.$$

The constant C does not depend on v. Differentiating both sides w.r.t. v and using $U'(v) = F(v)$ gives a first-order linear differential equation for β:

$$F(v) = F(v) + vf(v) - \beta(v)f(v) - \beta'(v)(1 - F(v)) + \beta(v)f(v),$$

where the $-\beta(v)f(v)$ and $+\beta(v)f(v)$ terms cancel, yielding

$$\beta'(v) = \frac{vf(v)}{1 - F(v)}.$$

Integrating from v^* to v and using $\beta(v^*) = \varepsilon$ yields (4).

Finally, cutoff indifference $U(v^*) = 0$ implies $v^*F(v^*) - \varepsilon = 0$, which is (3). □

A.4 Proof of Proposition 4.6

We compare the ERUS-optimal cutoff v^* with the standard revenue-optimal cutoff.

Lemma A.1 (ERUS-optimal cutoff condition)**.** *Suppose F is strictly monotone (Definition 4.5). The cutoff v^* that maximizes* ERUS(WOA(ε)) *is uniquely characterized by*

$$1 - F(v^*) = 2v^*f(v^*). \tag{11}$$

Proof. By Lemma 4.4, the equilibrium in WOA(ε) is parameterized by the cutoff v^* via $\varepsilon = v^*F(v^*)$. Under unlimited manipulation, bidder 1 pays her bid whenever she enters, so

$$\text{ERUS}(\text{WOA}(\varepsilon)) = \int_{v^*}^{1} \beta(v)\, dF(v).$$

Using (4) and Fubini's theorem,

$$\begin{aligned} ERUS &= \int_{v^*}^{1} \left(\varepsilon + \int_{v^*}^{v} \frac{tf(t)}{1 - F(t)}\, dt \right) dF(v) \\ &= (1 - F(v^*))\varepsilon + \int_{v^*}^{1} \left(\int_{v^*}^{v} \frac{tf(t)}{1 - F(t)}\, dt \right) f(v)\, dv \\ &= (1 - F(v^*))\varepsilon + \int_{v^*}^{1} \frac{tf(t)}{1 - F(t)} (1 - F(t))\, dt \\ &= (1 - F(v^*))\varepsilon + \int_{v^*}^{1} tf(t)\, dt. \end{aligned}$$

Substituting $\varepsilon = v^*F(v^*)$ gives

$$\text{ERUS}(v^*) = (1 - F(v^*))v^*F(v^*) + \int_{v^*}^{1} tf(t)\, dt.$$

Differentiating w.r.t. v^* yields

$$\frac{d}{dv^*}\text{ERUS} = F(v^*)(1 - F(v^*)) - 2v^*F(v^*)f(v^*) = F(v^*)\Big(1 - F(v^*) - 2v^*f(v^*)\Big).$$

For any interior cutoff $v^* \in (0, 1)$ we have $F(v^*) > 0$, so the first-order condition is

$$1 - F(v^*) = 2v^*f(v^*),$$

which is exactly (11).

Strict monotonicity of $g(v) = 1 - F(v) - 2vf(v)$ implies g crosses zero at most once, so the solution is unique. □

Proof of Proposition 4.6. Let v_{ERUS} solve the ERUS-optimal cutoff condition (11), i.e.,

$$g(v_{\text{ERUS}}) \equiv 1 - F(v_{\text{ERUS}}) - 2v_{\text{ERUS}}f(v_{\text{ERUS}}) = 0.$$

Let r_{ER} denote the standard revenue-optimal reserve (cutoff) in a truthful efficient auction, which solves the Myerson condition

$$1 - F(r_{\text{ER}}) = r_{\text{ER}}f(r_{\text{ER}}). \tag{12}$$

Evaluate g at r_{ER}:

$$g(r_{\text{ER}}) = 1 - F(r_{\text{ER}}) - 2r_{\text{ER}}f(r_{\text{ER}}) = r_{\text{ER}}f(r_{\text{ER}}) - 2r_{\text{ER}}f(r_{\text{ER}}) = -r_{\text{ER}}f(r_{\text{ER}}) < 0.$$

Under strict monotonicity (Definition 4.5), g is strictly decreasing and hence crosses zero at most once. Since $g(v_{\text{ERUS}}) = 0$ and $g(r_{\text{ER}}) < 0$, it follows that $v_{\text{ERUS}} < r_{\text{ER}}$.

Finally, in WOA(ε) the reserve parameter equals the bid threshold and satisfies $\varepsilon = vF(v)$ at the cutoff type v (equation (3)). Therefore

$$\varepsilon_{\text{ERUS}} = v_{\text{ERUS}}F(v_{\text{ERUS}}) \qquad \text{and} \qquad \varepsilon_{\text{ER}} = r_{\text{ER}}F(r_{\text{ER}})$$

(cf. (5) in the main text). Since $vF(v)$ is increasing on $[0, 1]$, $v_{\text{ERUS}} < r_{\text{ER}}$ implies $\varepsilon_{\text{ERUS}} < \varepsilon_{\text{ER}}$. □

A.5 Proof of Lemma 5.1

Proof. Fix bidder 3's bid $\beta_3 = 0$ and a non-delivery probability $q < \bar{q}(v_L) = 1 - 1/v_L$. Consider bidder 1 with budget c_1 and value v_L facing bidder 2 who bids $\beta_2 = c_2$ with $c_2 \sim \text{Unif}[0, 1]$.

If bidder 1 bids $\beta_1 \in [0, c_1]$, then in the delivery state she wins iff $c_2 < \beta_1$. In that event she obtains value v_L and pays c_2 (the highest losing bid). If she loses in the delivery state, she pays her own bid β_1. In the non-delivery state she also pays β_1.

Thus her expected payoff is

$$\begin{aligned} U_1(\beta_1) &= (1-q)\left(\int_0^{\beta_1}(v_L - c_2)\,dc_2 \;-\; \beta_1\int_{\beta_1}^1 dc_2\right) \;-\; q\beta_1 \\ &= (1-q)\left(v_L\beta_1 - \frac{\beta_1^2}{2} - \beta_1(1-\beta_1)\right) - q\beta_1 \\ &= \frac{1-q}{2}\beta_1^2 + ((1-q)v_L - 1)\beta_1. \end{aligned}$$

The derivative is $U_1'(\beta_1) = (1-q)\beta_1 + (1-q)v_L - 1$. Under $q < 1 - 1/v_L$ we have $(1-q)v_L - 1 > 0$, hence $U_1'(\beta_1) > 0$ for all $\beta_1 \in [0, 1]$. Therefore U_1 is increasing on $[0, c_1]$ and bidder 1's best response is $\beta_1 = c_1$. By symmetry, bidder 2's best response is $\beta_2 = c_2$. □

A.6 Proof of Proposition 5.2

Proof. Fix $q \in [\underline{q}(v_H), \bar{q}(v_L))$.

Step 1: bidders 1 and 2 best respond given $\beta_3 = 0$. By Lemma 5.1, if bidder 3 bids $\beta_3 = 0$ and $q < \bar{q}(v_L)$, then $\beta_1 = c_1$ and $\beta_2 = c_2$ are mutual best responses for bidders 1 and 2.

Step 2: bidder 3 best responds given $\beta_1 = c_1$ and $\beta_2 = c_2$. Now suppose bidders 1 and 2 bid $\beta_1 = c_1$ and $\beta_2 = c_2$ with $c_1, c_2 \sim \text{Unif}[0,1]$ independent, and consider bidder 3 with value v_H. Since opponents never bid above 1, bidding above 1 is weakly dominated for bidder 3 (it weakly increases delivery-state winning probability but strictly increases the non-delivery payment), so restrict attention to $\beta_3 \in [0,1]$.

If bidder 3 bids $\beta_3 \in [0,1]$, then in the delivery state he wins iff $c_1 < \beta_3$ and $c_2 < \beta_3$, which occurs with probability β_3^2. Thus his expected value term is $(1-q)\, v_H \beta_3^2$.

In the delivery state, conditional on winning, he pays the highest losing bid $M \equiv \max\{c_1, c_2\}$. Since M has density $2x$ on $[0,1]$ under i.i.d. uniforms, his expected payment from winning equals

$$(1-q)\int_0^{\beta_3} x \cdot 2x\, dx = \frac{2(1-q)}{3}\beta_3^3.$$

If he loses in the delivery state (probability $(1-q)(1-\beta_3^2)$), he pays his own bid β_3. In the non-delivery state (probability q), he also pays β_3. Therefore his expected payoff is

$$\begin{aligned} U_3(\beta_3) &= (1-q)\, v_H \beta_3^2 - \left(\frac{2(1-q)}{3}\beta_3^3 + (1-q)(1-\beta_3^2)\beta_3 + q\beta_3\right) \\ &= (1-q)\left(\frac{1}{3}\beta_3^3 + v_H\beta_3^2\right) - \beta_3. \end{aligned}$$

The second derivative is $U_3''(\beta_3) = 2(1-q)(\beta_3 + v_H) > 0$, so U_3 is strictly convex on $[0,1]$. Hence the maximum of U_3 over $[0,1]$ is attained at an endpoint: either $\beta_3 = 0$ or $\beta_3 = 1$. Since $U_3(0) = 0$, bidder 3's best response is $\beta_3 = 0$ iff $U_3(1) \le 0$. By (6), this holds iff $q \ge \underline{q}(v_H)$, which is true by assumption. Therefore bidder 3's best response is $\beta_3 = 0$.

Step 3: conclude equilibrium. Steps 1 and 2 show that the strategy profile stated in Proposition 5.2 is a Bayes–Nash equilibrium for $q \in [\underline{q}(v_H), \bar{q}(v_L))$. □

A.7 Proof of Proposition 5.3

Proof. Suppose, toward a contradiction, that $q \in (0, \underline{q}(v_H))$ and there exists a Bayes–Nash equilibrium with $\beta_1(c_1) = c_1$ and $\beta_2(c_2) = c_2$ almost surely.

Step 2 in the proof of Proposition 5.2 derives bidder 3's payoff $U_3(\beta_3)$ against opponents who bid their budgets, and shows that U_3 is strictly convex on $[0,1]$. Therefore any best response of bidder 3 lies in $\{0,1\}$. Since $U_3(0) = 0$ and

$$U_3(1) = (1-q)\left(v_H + \frac{1}{3}\right) - 1 > 0 \qquad \text{for } q < \underline{q}(v_H),$$

bidder 3's unique best response is to bid $\beta_3 = 1$.

Given $\beta_3 \equiv 1$, consider bidder $i \in \{1,2\}$ with any budget type $c_i > 0$. If she follows $\beta_i(c_i) = c_i$, she never wins (ties have probability zero) and pays c_i in both delivery and non-delivery states, yielding expected utility $-c_i$. By deviating to bid 0, she still never wins but pays 0 in all states, yielding utility $0 > -c_i$. Thus $\beta_i(c_i) = c_i$ is not a best response for any $c_i > 0$, contradicting the hypothesized equilibrium. □

A.8 Proof of Proposition 5.5

Proof. Consider $K \geq 2$ budget-constrained bidders with i.i.d. budgets $c_1, \dots, c_K \sim \text{Unif}[0,1]$ who bid $\beta_j = c_j$, and a deep-pocket bidder H with value v_H and bid $\beta \in [0,1]$.

Let $M = \max\{c_1, \dots, c_K\}$. Then $\mathbb{P}(M < \beta) = \beta^K$ and M has density Kx^{K-1} on $[0,1]$. In the delivery state, bidder H wins iff $M < \beta$; in that event he receives value v_H and pays M. Hence expected value from delivery is $(1-q)v_H\beta^K$ and expected *winning* payment is

$$(1-q)\mathbb{E}[M \cdot \mathbf{1}\{M < \beta\}] = (1-q)\int_0^\beta x \cdot Kx^{K-1}\,dx = \frac{K(1-q)}{K+1}\beta^{K+1}.$$

If bidder H loses in delivery, he pays his bid β; if non-delivery occurs, he also pays β. Therefore his expected utility is

$$\begin{aligned} U_H(\beta) &= (1-q)v_H\beta^K - \left(\frac{K(1-q)}{K+1}\beta^{K+1} + (1-q)(1-\beta^K)\beta + q\beta\right) \\ &= (1-q)\left(v_H\beta^K + \frac{1}{K+1}\beta^{K+1}\right) - \beta. \end{aligned}$$

The second derivative satisfies

$$U_H''(\beta) = (1-q)\Big(v_H K(K-1)\beta^{K-2} + K\beta^{K-1}\Big) = (1-q)K\beta^{K-2}(v_H(K-1)+\beta) \geq 0 \quad \text{for } \beta \in [0,1],$$

so U_H is convex on $[0,1]$ and the maximizer over $[0,1]$ is attained at an endpoint $\beta \in \{0,1\}$. In particular,

$$U_H(1) = (1-q)\left(v_H + \frac{1}{K+1}\right) - 1,$$

which yields the exit threshold in Proposition 5.5. □

A.9 Proof of Proposition 5.6

Proof. Fix any budget distribution G on $[0,1]$ and suppose bidders 1 and 2 bid their budgets, $\beta_i = c_i$, while bidder 3 bids 0. In the non-delivery state (probability q), no allocation occurs and total surplus equals minus total sunk costs, i.e., $-(c_1 + c_2)$. In the delivery state (probability $1-q$), the good is allocated to the higher-budget bidder, who receives value v_L. The winner pays the losing bid and the loser pays her own bid, so the total payment equals $2\min\{c_1, c_2\}$. Hence delivery-state surplus equals $v_L - 2\min\{c_1, c_2\}$.

Taking expectations yields

$$\begin{aligned} CS(q) &= (1-q)\mathbb{E}[v_L - 2\min\{c_1, c_2\}] + q\mathbb{E}[-(c_1 + c_2)] \\ &= (1-q)v_L - \Big(q\mathbb{E}[c_1 + c_2] + (1-q) \cdot 2\mathbb{E}[\min\{c_1, c_2\}]\Big). \end{aligned}$$

which is the expression stated in Proposition 5.6. □

A.10 Proof of Lemma 5.9

Proof. If bidders 1 and 2 obtain zero expected utility, total consumer surplus equals bidder 3's expected utility. Let $\omega \in \{0,1\}$ denote the non-delivery state, with $\Pr(\omega = 1) = q$. Bidder 3 obtains value only when delivery occurs ($\omega = 0$), and even then his winning probability is at most one. Moreover, payments are nonnegative in every state. Therefore bidder 3's expected utility is bounded by

$$U_3 \leq \mathbb{E}[\mathbf{1}\{\omega = 0\} \cdot v_H] = (1-q)v_H,$$

which implies the same bound for $\text{CS}(q)$. □

A.11 Proof of Proposition 5.10

Proof. At $q^* = \underline{q}(v_H)$, the exit-by-risk equilibrium is feasible by Proposition 5.2, and consumer surplus in that region is given by (7). Evaluating (7) at $q = q^*$ and using $1 - q^* = 1/(v_H + 1/3)$ yields

$$\text{CS}(q^*) = (1-q^*)v_L - \left(\tfrac{2}{3} + \tfrac{1}{3}q^*\right) = \frac{v_L + 1/3}{v_H + 1/3} - 1.$$

By Lemma 5.9, any dissipative equilibrium for $q < q^*$ satisfies $\limsup_{q \uparrow q^*} \text{CS}(q) \leq (1-q^*)v_H = \frac{v_H}{v_H + 1/3}$. Subtracting gives

$$\text{CS}(q^*) - (1-q^*)v_H = \frac{v_L - 2v_H}{v_H + 1/3},$$

which is strictly positive iff (8) holds. □

A.12 Proof of Corollary 5.12

Proof. Proposition 5.10 shows that $\text{CS}(q^*) - (1-q^*)v_H = \frac{v_L - 2v_H}{v_H + 1/3}$. If $v_L \leq 2v_H$, this expression is non-positive, yielding the stated inequality. □

A.13 Proof of Proposition 5.14

Proof. We prove each part of the proposition.

Part (i): Exit cutoff. Fix $q \in I$ and suppose the L-bidders bid their budgets, i.e., $\beta_i = c_i$ for $i \in \{1, \ldots, K\}$. Consider bidder H's payoff from bidding 1. Because all budgets lie in $[0,1]$, bidding above 1 is weakly dominated and bidding 1 guarantees that H is the (unique) highest bidder among all participants in the delivery state. Let $C_{(K)} \equiv \max\{c_1, \ldots, c_K\}$ denote the largest budget among the L-bidders.

If delivery occurs (probability $1-q$), bidder H receives value v_H and pays the highest losing bid, which equals $C_{(K)}$ because the L-bidders bid their budgets. Thus his delivery-state payoff is $v_H - C_{(K)}$. If non-delivery occurs (probability q), no allocation occurs and every bidder pays her own bid; in particular bidder H pays 1 and receives no value. Therefore

$$U_H(1) = (1-q)\,\mathbb{E}\big[v_H - C_{(K)}\big] - q \cdot 1 = (1-q)(v_H - \mu_{\max}) - q,$$

where $\mu_{\max} = \mathbb{E}[C_{(K)}]$. If $v_H > \mu_{\max}$, the equation $U_H(1) = 0$ has a unique solution

$$q^* = \frac{v_H - \mu_{\max}}{1 + v_H - \mu_{\max}} \in (0, 1).$$

Since $U_H(0) = 0$ and (by assumption (ii)) bidder H's best response against budget-bidding rivals is an endpoint $\beta_H \in \{0, 1\}$, we conclude that for any $q \geq q^*$ we have $U_H(1) \leq 0$ and hence H exits (bids 0). Combining this with assumption (i) (budget-bidding by the L-bidders is an equilibrium when H bids 0) implies that an exit-by-risk equilibrium exists for all $q \in I$ with $q \geq q^*$.

Part (ii): Consumer surplus in the exit equilibrium. Consider the exit-by-risk equilibrium in which H bids 0 and each L-bidder bids $\beta_i = c_i$. Let $C_{(1)} \leq \cdots \leq C_{(K)}$ denote the order statistics of $(c_1, \ldots, c_K)$.

In the non-delivery state (probability q), no allocation occurs and each bidder pays her bid. Hence total surplus equals minus total payments, i.e.,

$$\text{CS} = -\sum_{i=1}^{K} c_i.$$

In the delivery state (probability $1 - q$), the highest-budget bidder (with budget $C_{(K)}$) receives value v_L. The winner pays the highest losing bid $C_{(K-1)}$ and each of the $K - 1$ losers pays her own bid. Thus total payment in delivery equals

$$C_{(K-1)} + \sum_{j=1}^{K-1} C_{(j)} = \sum_{i=1}^{K} c_i \; - \; C_{(K)} \; + \; C_{(K-1)},$$

so delivery-state consumer surplus is

$$v_L - \Big(\sum_{i=1}^{K} c_i \; - \; C_{(K)} \; + \; C_{(K-1)}\Big) = v_L - \sum_{i=1}^{K} c_i + (C_{(K)} - C_{(K-1)}).$$

Taking expectations and using the definitions $\mu_{\text{sum}} = \mathbb{E}[\sum_{i=1}^{K} c_i]$ and $\mu_{\text{gap}} = \mathbb{E}[C_{(K)} - C_{(K-1)}]$ gives

$$\text{CS}_{\text{exit}}(q) = (1 - q)(v_L - \mu_{\text{sum}} + \mu_{\text{gap}}) + q(-\mu_{\text{sum}}) = (1 - q)v_L - \mu_{\text{sum}} + (1 - q)\mu_{\text{gap}}.$$

Part (iii): Jump condition relative to dissipative low-q equilibria. Fix any $q < q^*$ and consider a dissipative equilibrium in the sense of Definition 5.8. Lemma 5.9 implies the upper bound $\text{CS}(q) \leq (1 - q)v_H$.

Next, evaluate the exit-equilibrium surplus at $q = q^*$. Since $q^* = \frac{v_H - \mu_{\max}}{1 + v_H - \mu_{\max}}$, we have

$$1 - q^* = \frac{1}{1 + v_H - \mu_{\max}}. \tag{13}$$

Therefore, using Part (ii),

$$\text{CS}_{\text{exit}}(q^*) > (1 - q^*)v_H \iff (1 - q^*)(v_L - v_H + \mu_{\text{gap}}) > \mu_{\text{sum}} \iff v_L - v_H + \mu_{\text{gap}} > \mu_{\text{sum}} \frac{1}{1 - q^*}$$

and substituting (13) yields

$$\mathrm{CS}_{\mathrm{exit}}(q^*) > (1-q^*)v_H \iff v_L - v_H + \mu_{\mathrm{gap}} > \mu_{\mathrm{sum}}(1 + v_H - \mu_{\max}),$$

which is the stated condition. If the inequality fails weakly, then $\mathrm{CS}_{\mathrm{exit}}(q^*) \le (1-q^*)v_H$, so the exit equilibrium does not exceed the dissipative upper bound at the cutoff. □

B A robustness extension for Section 4

This short appendix note shows that the mechanism ranking in Theorem 4.2 is unchanged when seller-side cancellation is probabilistic or costly.

Proposition B.1 (Probabilistic or costly cancellation)**.** *Fix a benchmark mechanism M and suppose bidders remain naïve about cancellation (they bid as in the perceived game with $\hat{q} = 0$). After bids are submitted, the platform can cancel the auction (trigger non-delivery) with probability $\lambda \in [0, 1]$, chosen ex ante, and incurs a cost $k(\lambda)$. Conditional on cancellation, the platform can implement the benchmark manipulation from (1). Let $R(M)$ denote the platform's expected revenue in the delivery state under honest play of the advertised mechanism M. Then the platform's total expected payoff equals $(1-\lambda)R(M) + \lambda \cdot \mathrm{ERUS}(M) - k(\lambda)$. Since $R(M)$ is the same across admissible efficient auctions by revenue equivalence, the mechanism ranking for any fixed λ and any cost function k is governed by* $\mathrm{ERUS}(M)$*. Accordingly, holding fixed the delivery-state revenue, the platform's net payoff from cancellation states equals*

$$\lambda \cdot \mathrm{ERUS}(M) - k(\lambda).$$

Hence $\mathrm{WOA}(0)$ *remains optimal among admissible efficient auctions.*

Proof. Because bidders ignore cancellation, the equilibrium bidding function in the advertised mechanism M is independent of λ. Conditional on cancellation, the payment collected from the targeted bidder coincides with the payment in (1); thus expected cancellation revenue is $\lambda\mathrm{ERUS}(M)$. Subtracting the cost $k(\lambda)$ yields the stated expression. For any fixed λ, the term $-k(\lambda)$ does not depend on the mechanism, so the mechanism ranking is determined by $\mathrm{ERUS}(M)$. The final claim follows from Theorem 4.2. □

C Supplementary results: endogenous participation in the L-subgame

This appendix note formalizes a convenient reduced form for the "L-only" subgame (when the deep-pocket bidder exits). In many sunk-cost contests, a high-value bidder with a very small budget may prefer to abstain (bid zero) rather than incur a payment that is likely to be wasted. Our main text abstracts from the population-to-participant selection map by treating the induced distribution of *effective bids* among active L-bidders as primitive. The results below justify this practice by providing a general payoff expression and a simple sufficient condition for *bang–bang* best responses (bid either 0 or the full budget).

We follow Section 3 in adopting a *strict* reserve rule: bids at the reserve are treated as non-entry. In particular, in $\mathrm{WOA}(0)$ bidding 0 is inactive, yields zero payoff, and cannot win.

Lemma C.1 (Bang–bang best responses under convexity)**.** *Fix* $K \geq 2$ *and consider the* L*-subgame in* WOA(0) *under the aware regime* $\hat{q} = q$*. Fix a CDF* F *on* $[0, 1]$ *that describes the distribution of a rival's* active *bid and satisfies* $F(0) = 0$*. For a bidder with value* v_L *and budget* c*, define for* $b \in [0, c]$ *the interim payoff*

$$u(b) \equiv -b + (1 - q)\Big(v_L F(b)^{K-1} + \int_0^b F(x)^{K-1}\, dx\Big). \tag{14}$$

If the function $b \mapsto F(b)^{K-1}$ *is convex on* $[0, 1]$*, then* $u(\cdot)$ *is convex on* $[0, 1]$*. Consequently, for every budget type* c*, there exists an optimal bid in* $\{0, c\}$*.*

Proof. Let $h(b) \equiv F(b)^{K-1}$. By assumption, h is convex and nondecreasing on $[0, 1]$ (since F is a CDF). The map $b \mapsto \int_0^b h(x)\, dx$ is therefore convex as well. Equation (14) shows that $u(b)$ is the sum of a linear function ($-b$) and a positive multiple of convex functions, hence u is convex. Maximizing a convex function over a compact interval $[0, c]$ attains its maximum at an endpoint, so an optimal bid exists in $\{0, c\}$. □

Remark C.2 (Scope of the no-atom-at-0 claim)**.** The restriction $F(0) = 0$ in Lemma C.1 is intended for *active positive-budget* L*-types* in the symmetric L-subgame under the strict-reserve convention. If such types placed an atom at 0 and the all-zero event had positive probability, then deviating from 0 to a sufficiently small *active* bid $b > 0$ would win on that event while paying an arbitrarily small expected cost. Hence, in continuous bid spaces, no equilibrium of that L-subgame can assign positive mass at 0 to active positive-budget types when the all-zero event has positive probability. This is not a global prohibition on every player bidding 0 in the full game; in particular, it does not rule out bidder 3's pure exit ($\beta_3 = 0$) in Proposition 5.2.