MARKET-PRICE LEARNING UNDER ENDOGENOUS BID CENSORING: LIMITS, IDENTIFICATION, AND CORRECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

In real-time bidding, a bidder's shading and pacing decisions depend on the distribution of the market price it must beat. The bidder learns this distribution from its own auction logs, which record every bid but reveal the price only when the bid wins. Statistically, this is estimation under dependent censoring in which the censoring variable, the bid, is always observed. Standard censored estimation relies on bid–price independence, but bids respond to the signals that drive prices, making the censoring *endogenous*. Without any assumption on this dependence, the logs identify the price margin only up to a sharp band whose width is the share of auctions lost with lower bids. This ambiguity yields an assumption-free minimax error floor. Under affiliation between the bid and the price, the band narrows to a closed-form interval, and standard censored estimation understates the price CDF. To correct this bias, we assume a concordance-ordered copula family. Given this family, the observed bids and win-region cells identify the dependence parameter and the price distribution through the largest bid. The sharp band then collapses to a point and the minimax floor vanishes over the same range. We fit neural bid and price distributions in two stages, first from the fully observed bids, then through a Gaussian-copula observed-data likelihood. For the finite model under standard regularity conditions, the two-stage estimator is consistent and asymptotically normal at the root-n rate. On AuctionNet the correction cuts the largest gap of the censored network's CDF from 0.14 to 0.04 and closes about three quarters of its CRPS gap to the oracle. The gain persists when the bidding rule uses a market signal withheld from every estimator and when the coupling is planted outside the fitted family. It also reaches the decision layer, where the corrected bids cut the overshoot of a 10% target win rate on AuctionNet from 17.8 to 7.9 pp.

1 INTRODUCTION

In online advertising, a demand-side platform bids for impressions on behalf of advertisers. How much to shade a bid, how fast to spend a budget, and what surplus a campaign can expect all rest on the market-price distribution (Zhang et al., 2014a; Xu et al., 2015; Gligorijevic et al., 2020). The platform learns that distribution from its own auction logs.

For every auction it enters, the bidder observes one bit and, sometimes, one number. An opportunity with features X arrives. The bidder submits a bid B and the market settles at a price Z. Winning reveals Z exactly when $Z \leq B$. Losing reveals only $Z > B$, while the bid remains observed. These are the outcomes of a second-price auction, or of a first-price auction returning the minimum bid to win. The bidder's action thus decides which prices enter its training data (Achddou et al., 2021).

Existing bid-landscape models treat losses as right-censored observations and maximize the censored likelihood (Wu et al., 2015; Ren et al., 2019; Zhou et al., 2021; Li et al., 2022). Their consistency relies on the bid being independent of the price given X, which a signal outside X can break by moving both. Such signals are common. Bidders value the impression from the same request and their own data, only part of which the price model retains in X. Shading from a price forecast moves the bid with the price it predicts. Pacing a budget raises and lowers the bid by the time of day, which also moves the price. The censoring is then *endogenous*. On AuctionNet (Su et al., 2024), whose

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

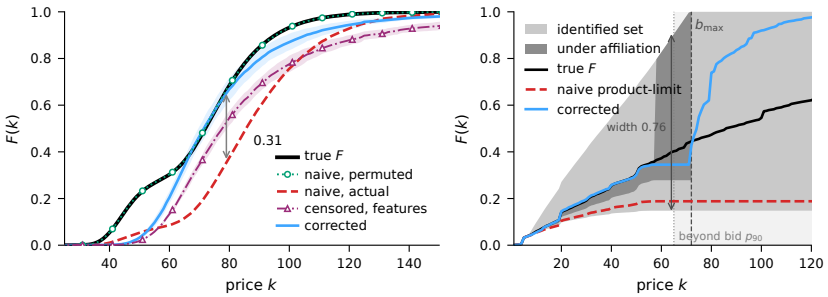

Figure 1: Endogenous bid-censoring and the sharp identified set. (Left) On AuctionNet the naive product-limit CDF misses the truth by 0.31. Permuting bids makes censoring independent and reduces the error to 0.004. A censored network with features misses by 0.14. Our correction reduces the gap at that price to 0.04. (Right) The sharp identified set of Proposition 1 on iPinYou at coupling $\lambda = 0.5$. The corrected curve flattens above the largest winning price, 57. The arrow marks the width 0.76 at the bid p_{90} (dotted line), whose half 0.38 is the minimax floor of Corollary 1. The darker band holds under affiliation (Theorem 1(i)). Theorem 1 ends at the largest bid (dashed line).

agents bid from the signals that also set the price, the product-limit estimate (Kaplan & Meier, 1958) misses the marginal price CDF by up to 0.31. Permuting the bids across auctions and rebuilding the censored records keeps both margins and cuts the same error to 0.004. A censored network with the observed features still misses by 0.14 (Figure 1, left).

We first ask what the censored log reveals without assumptions on the dependence between Z and B. The marginal price distribution is then identified only up to a sharp band, whose width at price k is the probability of losing with a bid below k (Proposition 1). Two bid–price distributions at opposite boundaries generate the same observable data, so no amount of it can tell them apart. The resulting ambiguity yields a minimax error floor that no estimator can overcome (Corollary 1). This band is the classical dependent-censoring bound adapted to auctions (Tsiatis, 1975; Peterson, 1976).

Independence closes this gap, making standard censored estimation consistent. Auction theory, however, models such shared value signals as affiliated, so a higher bid tends to come with a higher price (Milgrom & Weber, 1982; Athey & Haile, 2002; Krishna, 2010). Under affiliation, Proposition 2 shows that standard censored estimation understates the price CDF. Proposition A.1 traces this downward bias to two factors, how many auctions were already lost with a lower bid and how far their hazard is from the observed one. The bias vanishes at independence and grows as dependence strengthens under the condition of Proposition A.2. The same assumption also narrows the band to sharp closed-form bounds (Theorem 1(i)).

To address this bias, we model bid–price dependence with a concordance-ordered copula family. This assumption alone does not identify the model under standard right censoring, where the censoring time is as hidden as the event (Zheng & Klein, 1995; Czado & Van Keilegom, 2023). In an auction, by contrast, the bid is on record whether it wins or loses. The won auctions then determine one dependence parameter and one price margin within any concordance-ordered family (Theorem 1(ii)). Consequently, the sharp band of Proposition 1 shrinks to a point through the largest bid (Corollary 2), taking the floor of Corollary 1 with it.

This result motivates a two-stage estimator with neural margins. The first stage fits the fully observed bid distribution. The second fits the price margin and the Gaussian copula parameter by maximizing the observed-data likelihood. Because that likelihood uses the exact cell probabilities of the family, Theorem 1 identifies the shared dependence parameter and the conditional margins of the fitted model. Theorem 2 gives consistency and root-n normality under standard regularity conditions.

Our main contribution is to characterize and measure the bias that residual bid–price dependence leaves in market-price learning and to assess its correction. Across three benchmarks the correction removes most of the censoring bias, closing about three quarters of the censored network's CRPS gap to the oracle on AuctionNet, and keeping the gain under a withheld price signal or a coupling outside the Gaussian family (Table 2). The gain reaches decisions as well. On iPinYou the correction cuts

the censored network's mean win-probability error from 6.1 to 0.9 pp. In bid replay on AuctionNet, the corrected bids cut the overshoot of a 10% target win rate from 17.8 to 7.9 pp (Table 3).

2 PROBLEM SETUP AND THEORY

Notation. Prices live on the grid $[K] = \{1, \dots, K\}$. Each auction opportunity has features X, a market price $Z \in [K]$, and a bid $B \in [K]$. The bidder observes (X, B, W, WZ) with $W = \mathbf{1}\{Z \le B\}$. As in second-price auctions and their relatives, a win reveals the price exactly while a loss reveals only $Z > B$.[1] We write $F(k) = \mathbb{P}(Z \le k)$ and $G(k) = \mathbb{P}(B \le k)$ for the price and bid margins, with pmfs $f(k) = \mathbb{P}(Z = k)$ and $g(k) = \mathbb{P}(B = k)$. Set $F(0) = G(0) = 0$. We denote the joint law by $\pi(z, b) = \mathbb{P}(Z = z, B = b)$ and the per-bid loss mass by $\ell(b) = \mathbb{P}(W = 0,\ B = b)$. Throughout the paper, *margin* always means a marginal distribution. The win region is $\{(z, b) : z \le b\}$. We suppress conditioning on X until the conditional results below.

2.1 WHAT THE CENSORED AUCTION LOG CAN AND CANNOT IDENTIFY

Without any assumption on the bid–price dependence, the observable law determines the price margin only up to a band.

Proposition 1 (Sharp identified set)**.** *Two joint laws of (Z, B) are observationally equivalent if and only if they share the bid margin and the restriction of π to the win region. Consequently, for every $k < K$ the identified set for $F(k)$ is the interval*

$$\Big[\ \mathbb{P}(W{=}1,\ Z \le k),\ \ \mathbb{P}(W{=}1,\ Z \le k) + \mathbb{P}(W{=}0,\ B < k)\ \Big],$$

and both endpoints are attained. The ambiguity at price k is therefore exactly the probability of losing with a bid below k.

To see the bound, split $F(k)$ into its win and loss parts. A loss with $B \ge k$ must have $Z > k$, whereas a loss with $B < k$ may fall on either side of k.

Proposition 1 is the bid-censoring form of the classical bounds under dependent censoring (Tsiatis, 1975; Peterson, 1976). Auction identification has borrowed such arguments before. Athey & Haile (2002) use them to recover the distribution of bidders' valuations. There each valuation is censored by the rival valuations, which are themselves latent, whereas here the censoring variable is the always-observed bid, so the ambiguity width at k is itself observable.

Figure 1 (right) shows the band on the iPinYou benchmark. The naive product-limit stays near the lower endpoint of the set once losses dominate.

Corollary 1 (Assumption-free minimax floor)**.** *Let P and P' be observationally equivalent laws whose margins differ by $\Delta = \|F - F'\|_\infty$. Then every estimator $\hat{F}$ of the margin satisfies, at every sample size,*

$$\max_{Q \in \{P, P'\}} \mathbb{E}_Q\Big[\|\hat{F} - F_Q\|_\infty\Big] \ge \Delta/2,$$

where F_Q is the price margin under Q.

2.2 THE DIRECTION OF THE NAIVE BIAS

Proposition 1 bounds where the truth can live. The next question is which point inside the band naive estimation delivers. Let $h(k) = \mathbb{P}(Z{=}k \mid Z \ge k)$ be the true hazard, $h^{\#}(k) = \mathbb{P}(Z{=}k \mid Z \ge k,\ B \ge k)$ the crude hazard that censored estimation actually averages, and $F^{\#}(k) = 1 - \prod_{j \le k}(1 - h^{\#}(j))$ the population limit of the product-limit (Kaplan–Meier) estimator. The product-limit is the nonparametric maximizer of the censored likelihood. It builds the CDF from hazards, which remain estimable from the at-risk set under censoring. Dependence therefore enters one hazard at a time.

Proposition 2 (Affiliation implies downward bias)**.** *Suppose the conditional hazards $q_b(k) = \mathbb{P}(Z{=}k \mid Z \ge k,\ B{=}b)$ are nonincreasing in b for each k, as holds when (Z, B) are affiliated in the sense of Milgrom & Weber (1982). Then $h^{\#}(k) \le h(k)$, and therefore $F^{\#}(k) \le F(k)$, at every k with $\mathbb{P}(Z \ge k,\ B \ge k) > 0$. Naive censored estimation* understates *the price distribution.*

[1] A strict win rule $Z < B$ reduces to this convention by relabeling $b \mapsto b - 1$.

The monotone-hazard premise is strong, and affiliation is the standard condition that delivers it. Under affiliation, a higher bid shifts the conditional price distribution upward in the likelihood-ratio order, and that order lowers the hazard at every k. Conditioning on features removes this bias only if it removes the dependence itself. Figure 1 (left) shows how much of the gap features close.

The direction settled, what remains is the size of the bias and the factors that move it. For every k with $\mathbb{P}(Z \geq k,\, B \geq k) > 0$ and $\mathbb{P}(Z \geq k,\, B < k) > 0$, the two hazards obey the exact identity

$$h(k) - h^{\#}(k) \;=\; \alpha(k)\,\big(A(k) - h^{\#}(k)\big),$$

where $\alpha(k) = \mathbb{P}(B < k \mid Z \geq k)$ is the share of the surviving price mass held by already-lost auctions and $A(k) = \mathbb{P}(Z{=}k \mid Z \geq k,\, B < k)$ is their hazard. The hazard gaps accumulate to $|F(k) - F^{\#}(k)| \leq \sum_{j \leq k} |h(j) - h^{\#}(j)|$ (Proposition A.1, Appendix A).

The bias is also monotone in the *strength* of the (Z, B) dependence. Under the copula families of Proposition A.2 (Appendix A), the gap $F - F^{\#}$ is zero at independence and never shrinks as the dependence strengthens. Appendix A checks its condition on our benchmark margins.

2.3 What additional structure identifies

Recovering F inside the band requires an assumption on the dependence. One remedy is exogenous bid variation, bid exploration in deployed bidders that shifts only the bid (Huang et al., 2026). A bidding system whose bids decompose as $B = \mu(X, \varepsilon)$ with pacing noise $\varepsilon \perp Z \mid X$ is one example. When such variation accounts for the whole bid given X, so that $Z \perp B \mid X$, the crude hazard coincides with the true hazard within every cell. The product-limit then recovers $F(\cdot \mid x)$ up to the largest bid in the conditional support. The survival-analysis counterpart pairs exogenous variation in a confounded covariate with a parametric joint model (Crommen et al., 2024). We instead assume that the joint law of (Z, B) lies in a specified copula family. This assumption is of the same kind as independence but weaker, since the family includes independence.

Copulas. By Sklar's theorem (Sklar, 1959), any joint law of $Z \sim F$ and $B \sim G$ can be written as $\mathbb{P}(Z \leq z,\, B \leq b) = C(F(z), G(b))$ for a copula C. We assume that C belongs to a one-parameter family $\{C_\theta : \theta \in \Theta\}$ ordered by concordance, so that increasing θ raises $C_\theta(u, v)$ at every (u, v) and the parameter measures the strength of dependence. Copula survival analysis makes a similar assumption, fixing the copula itself (Zheng & Klein, 1995) or a family with a free association parameter (Czado & Van Keilegom, 2023; Deresa & Van Keilegom, 2021; Foomani et al., 2023), and still needs further restrictions on the margins to identify the model. Here the censoring variable is observed. Theorem 1 states what each assumption yields in turn. Affiliation alone narrows the band to a sharp one, and a copula family picks out one point inside it.

Theorem 1 (Identification under affiliation, and within a family)**.** *Let b_0 and $b_{\max}$ be the smallest and largest support points of the bid.*

(i) *Suppose (Z, B) are affiliated, so that $\pi(z, b)\,\pi(z', b') \geq \pi(z, b')\,\pi(z', b)$ whenever $z < z'$ and $b < b'$. Then $F_L(k) \leq F(k) \leq F_U(k)$ at every $k \leq b_{\max}$, where F_L and F_U are functions of the observed cells given in Appendix A.3. Both bounds are sharp, each attained by one affiliated law at every k at once.*

(ii) *Fix an ordered parameter space Θ and a copula family $\{C_\theta : \theta \in \Theta\}$ for which $\theta \mapsto C_\theta(u, v)$ is strictly increasing at every $(u, v) \in (0, 1)^2$. Suppose each C_θ is absolutely continuous with density c_θ strictly positive on $(0, 1)^2$. Couple $Z \sim F$ and $B \sim G$ through C_θ, with $b_{\max} > b_0$ and $0 < F(b_0) < 1$. Then the observed bid margin and the win-region cells uniquely determine θ and $F(k)$ for all $k \leq b_{\max}$. The recovery is sequential.*

The lower edge of part (i) places each lost price as high as affiliation allows, and the upper edge places it as low. For the Gaussian with $\rho \geq 0$, the point of part (ii) lies inside this band. On the iPinYou margins, the band under affiliation averages 0.02 to 0.07 in width, against 0.33 without assumptions (Appendix B.7). The Gaussian family of Section 3 meets both hypotheses of part (ii), with a strictly positive density and the concordance ordering shown by Plackett (1954).

The proof of part (ii) recovers θ and F step by step. The first step uses the prices at or below b_0, where the price never exceeds the bid and every auction is won. These cells determine $F(b_0)$.

The cell with bid b_0 then determines θ through a strictly increasing equation. Once θ is known, each larger win-region cell gives one equation for the next value of F, a recursion that positive copula density keeps one-to-one through $b_{\max}$. This is also a sequential estimator, the copula-graphic recursion (Zheng & Klein, 1995) with the bid margin fully observed rather than recovered. Appendix A.3 examines identification when no price falls at or below the smallest bid.

Corollary 2 (Point identification through the largest bid)**.** *Under the hypotheses of Theorem 1(ii), two joint laws in the same copula family are observationally equivalent if and only if they share the bid margin, the dependence parameter, and $F(k)$ at every $k \le b_{\max}$. Within the family, $F(k)$ is therefore point identified for every $k \le b_{\max}$. The interval of Proposition 1, of width $\mathbb{P}(W{=}0,\, B < k)$, collapses to that point there. The floor of Corollary 1 vanishes with it on the bid support. For $b_{\max} < k < K$ the identified set is $[F(b_{\max}),\, 1]$.*

Part (ii) rests on two things the log cannot check, the anchor condition $0 < F(b_0) < 1$ and the family itself. Because the bid is the bidder's own action, a small share of auctions can be held at a fixed bid t, which strengthens the theorem on both counts. The anchor moves to t, and the marked share yields a margin obtained without the family, against which a wrong family is rejected (Proposition A.3). In simulation, a 5% share recovers F above t to about 0.002 on a Gaussian stream, with or without the anchor, and rejects a wrong family in every draw (Appendix B.6).

2.4 What the estimator guarantees

We state the guarantee without features first and return to the proposed model at the end of this subsection. Stage one estimates the bid pmf by empirical frequencies $\hat{g}$, with $\hat{G}(k) = \sum_{j \le k} \hat{g}(j)$. Stage two maximizes the exact observed-data log-likelihood

$$L_n(\theta, F) = \sum_{z \le b} n_{zb}\, \log \pi_{\theta, F, \hat{G}}(z, b) + \sum_{b} n_b^{\text{loss}}\, \log \ell_{\theta, F, \hat{G}}(b)$$

over $\Theta \times \mathcal{F}_\delta$, where n_{zb} and n_b^{loss} count win and loss cells, π and ℓ are their copula rectangle probabilities, and $\mathcal{F}_\delta$ constrains the hazards of F on $\{1, \dots, b_{\max}\}$ to $[\delta, 1{-}\delta]$ for fixed $\delta \in (0, 1/2)$.

The guarantee needs regularity. The auctions are i.i.d. from a member of a family satisfying the hypotheses of Theorem 1(ii), with $b_{\max} < K$. The true parameter θ_0 is interior to Θ and the true hazards lie in $(\delta, 1-\delta)$. The modeled cell probabilities are smooth, bounded away from zero, and have a nonsingular expected Hessian at the truth. Appendix A states these as conditions (R1)–(R4).

Theorem 2 (Consistency and root-n normality)**.** *Write*

$$\eta = \big(\theta, F(1), \dots, F(b_{\max})\big)$$

and define $\hat{\eta}$ and η_0 in the same way. Under (R1)–(R4), $\hat{\eta} \to \eta_0$ in probability and

$$\sqrt{n}(\hat{\eta} - \eta_0) \;\xrightarrow{d}\; \mathcal{N}\big(0,\, \mathcal{I}^{-1} \Omega \mathcal{I}^{-1}\big),$$

where Ω is the variance of the stage-two score corrected for the influence of the empirical bid margin. Both matrices are given in Appendix A.

The sandwich covariance accounts for the first-stage estimation of the bid margin and yields asymptotically valid confidence intervals for θ and $F(k)$ from the censored log alone. Identification, which any likelihood guarantee assumes, comes from Theorem 1. Nonsingularity of $\mathcal{I}$, condition (R4), is the one hypothesis the structure does not supply. It holds in the well-specified check of Appendix B.5, where the sandwich intervals keep their coverage. The proposed model of Section 3 runs this estimator at each feature value with one shared θ. Over finitely many feature values, each recurring often enough to fill its cells, the theorem carries over to the saturated fit (Appendix A.5).

3 Two-stage Gaussian-copula censored likelihood

Our estimator couples neural bid and price margins through a Gaussian copula with a shared dependence parameter and fits them in two stages. Stage one fits the bid margin $G(\cdot \mid x)$, which is fully observed. Stage two fixes $\hat{G}$ and maximizes the conditional observed-data objective

$$(\hat{F}, \hat{\theta}) = \arg\max_{F, \theta} \sum_{i \text{ won}} \log\big(F_{c,\hat{G}}(z_i \mid x_i, b_i) - F_{c,\hat{G}}(z_i - 1 \mid x_i, b_i)\big) + \sum_{i \text{ lost}} \log\big(1 - F_{c,\hat{G}}(b_i \mid x_i, b_i)\big),$$

where $F_{c,\hat{G}}$ is the price CDF conditional on the features and the bid, built by coupling F and $\hat{G}$ through the Gaussian copula C_θ of Section 2.

A bid backbone with parameters ϕ solves

$$\hat{\phi} = \arg\max_{\phi} \sum_i \log[G_\phi(b_i \mid x_i) - G_\phi(b_i - 1 \mid x_i)], \qquad \hat{G}(\cdot \mid x) = G_{\hat{\phi}}(\cdot \mid x).$$

A price backbone with parameters ψ outputs discrete hazards,

$$h_{\psi,k}(x) = \sigma(a_{\psi,k}(x)), \qquad F_\psi(k \mid x) = 1 - \prod_{j \le k} \big(1 - h_{\psi,j}(x)\big),$$

with σ the logistic sigmoid. The Gaussian copula couples the two margins on the bid cell. For bid b, the conditional CDF is the copula mass of the bid cell below price k divided by the cell's mass,

$$\begin{aligned} F_{c,\hat{G}}(k \mid x, b) &:= \Pr_{\theta, F_\psi, \hat{G}}(Z \le k \mid X = x, B = b) \\ &= \frac{C_\theta(F_\psi(k \mid x), \hat{G}(b \mid x)) - C_\theta(F_\psi(k \mid x), \hat{G}(b-1 \mid x))}{\hat{G}(b \mid x) - \hat{G}(b-1 \mid x)}. \end{aligned}$$

Summed over b, the numerator telescopes to $F_\psi(k \mid x)$, so the price margin of the fitted joint law is the network output itself. At $\theta = 0$ the conditional CDF reduces to $F_\psi(k \mid x)$ and the objective to the standard censored likelihood, the ablation of Section 4. Stage two substitutes this conditional CDF into the objective and optimizes (ψ, θ), with $\theta = 0.99 \tanh(t)$ for a free scalar t; a feature-dependent t does not help (Appendix B.3). Because the first stage is fixed, this objective has the same maximizer as the joint likelihood L_n of Theorem 2 (Appendix B.1). Appendix B gives details.

4 Experiments

Data. AuctionNet is a calibrated synthetic benchmark with feature-rich impressions and ground-truth prices (Su et al., 2024). Its bidding agents set their bids from the signals that also move the price. The bid–price dependence is therefore the simulator's own. We keep its bid–price pairing. Its price is the least winning cost of a multi-slot auction, which the payment need not equal. The iPinYou and AuctionGym benchmarks keep the true price while sweeping the bid–price coupling λ from independence to strong alignment at fixed margins (Zhang et al., 2014b; Ren et al., 2019; Jeunen et al., 2023). iPinYou logs second-price auctions, so its price is also the payment. AuctionNet and iPinYou evaluate the proposed model. AuctionGym evaluates the estimator without features.

Controlled dependence. A further AuctionNet experiment simulates a bidder who acts on a signal the estimator never sees. We withhold the time step, which explains half of the price variance, from every estimator. A synthetic rule then sets the bids by mixing it, at weight $\gamma \in \{0, 0.5, 1\}$, with the visible features. Four constructions on iPinYou and AuctionGym plant the dependence outside the fitted family, through Clayton and Gumbel copulas, affiliated signals (Milgrom & Weber, 1982), and shading from a price forecast (Zhang et al., 2014a; Gligorijevic et al., 2020) (Table 2, Appendix B.4).

Baselines. We group the comparisons by how the fit treats the bid–price dependence. Three estimators assume independence. The naive product-limit uses no features. The censored network runs the proposed model's backbone and features under the standard censored likelihood. DLF replaces that backbone with a recurrence over the price grid (Ren et al., 2019). The covariate network takes the bid as an input under the same likelihood. Two estimators learn the dependence. The proposed model fits the full auction log. DCSurvival (Zhang et al., 2024) fits a neural copula to the survival record alone, which drops the bid on wins. Two further comparisons set the ceiling. The copula-graphic recursion of Theorem 1 solves the observed cells with the true dependence supplied. The oracle network is trained on the benchmark's uncensored prices. Neural results are averaged over ten seeds, with standard deviations. Appendix B.2 gives the data, splits, and training details.

Bid–price dependence biases the censored fit. On AuctionNet, the censored network's CDF averages 0.097 below the truth over the bid range, against 0.006 for the oracle network, the direction Proposition 2 predicts (Table 1). Without the true price, the log alone puts the product-limit below

Table 1: Margin recovery. CDF error is the mean absolute gap between the fitted and true marginal CDF over the bid range; signed CDF error keeps the sign, so a negative value is a downward shift. Column construction in Appendix B.

Benchmark	Metric	Independence*	Covariate	Corrected	Copula-graphic	Oracle
AuctionGym ($\lambda = 0.75$)	CDF error	0.081	–	0.022±0.000	0.021	–
AuctionGym ($\lambda = 0$)	CDF error	0.001	–	0.002±0.000	0.001	–
iPinYou ($\lambda = 0.5$)	CDF error	0.058±0.006	0.035±0.006	0.009±0.003	0.007	0.004±0.003
AuctionNet	CRPS	0.052±0.004	0.053±0.004	0.042±0.002	–	0.039±0.001
AuctionNet	signed CDF error	−0.097±0.012	–	−0.044±0.019	–	−0.006±0.002

*Independence denotes the estimator that assumes bid–price independence, the censored network on the feature rows and the naive product-limit on the AuctionGym rows, which have no features.

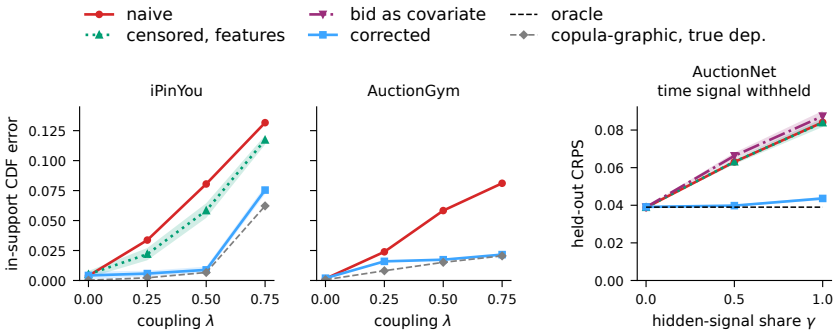

Figure 2: In the left and center panels the naive and censored errors rise with the coupling while the corrected error stays below 0.02 through $\lambda = 0.5$. (Left) CDF error within the bid range against the coupling λ on iPinYou. (Center) AuctionGym, where the corrected curve is the estimator without features. The copula-graphic curve runs the recursion with the true dependence. (Right) Held-out CRPS on AuctionNet against the withheld-signal share γ. Bands cover one standard deviation.

the lower edge of Theorem 1(i), and hence below the truth under affiliation (Appendix B.7). Along the λ grid of Figure 2, the naive and censored errors grow with λ on iPinYou and AuctionGym, in the direction Proposition A.2 predicts, while staying within 0.005 of the truth at independence. The same holds for the withheld signal. As the bidding rule's share γ of it grows from 0 to 1, the censored network's CRPS rises from 0.039 to 0.084, no better than the product-limit (Figure 2, right). DLF inherits the bias, 0.060 at $\lambda = 0.5$ against 0.058 for the censored network. The failure follows the objective rather than the architecture (Table 4).

The copula correction closes most of the gap to the oracle. On AuctionGym at $\lambda = 0.75$, the CDF error falls from 0.081 to 0.022, within 0.001 of the copula-graphic recursion given the true dependence (Table 1). On AuctionNet the censored network's CRPS of 0.052 sits 0.013 above the oracle's 0.039. The proposed model brings it to 0.042, removing 75% of that gap. On iPinYou at $\lambda = 0.5$, the rank mixture matches no single family, yet the CDF error falls from 0.058 to 0.009 against 0.004 for the oracle, within 0.002 of the copula-graphic recursion given the true dependence and far below the recursion under independence (Table 4). With the time step withheld, the model stays within 0.005 of the oracle at every γ (Figure 2, right). Its fitted dependence $\hat{\theta}$ vanishes at $\gamma = 0$ and turns positive for $\gamma > 0$ (Appendix B.4).

Fixing the dependence at zero turns the cell likelihood back into the censored likelihood. The fit then stays within 0.003 of the censored network on AuctionNet and iPinYou, so most of the gain comes from the estimated dependence. Learning it lowers the AuctionNet CRPS from 0.049 to 0.042 and the iPinYou CDF error from 0.059 to 0.009, closing 70% of the zero-dependence fit's gap to the oracle on AuctionNet. The correction needs the dependence near the truth. Given half the true correlation, the copula-graphic recursion on iPinYou improves only from 0.074 to 0.052 (Table 4).

At $\lambda = 0.75$ on iPinYou, the corrected error is still 0.075. The copula-graphic recursion given the true dependence degrades with λ in the same way, so the loss is not in the dependence estimate

Table 2: CDF error under couplings outside the fitted family, with the bid margin unchanged. Planted rows couple iPinYou with a Clayton or Gumbel copula; signal rows let a bidder's signal create the dependence. The corrected estimator uses no features, with the Gaussian family fixed or chosen by likelihood on the planted rows. The recursion receives the true Kendall dependence (Appendix B.4).

Coupling	**Naive**	**Corrected**		**Recursion**
		Gaussian	selected	true dependence
Clayton, lower tail	0.096	0.024	0.0003	0.002
Gumbel, upper tail	0.075	0.016	0.0007	0.002
Affiliated signals, AuctionGym ($\tau = 0.36$)	0.088	0.023	–	0.020
Predictor bidder, iPinYou ($\tau = 0.47$)	0.084	0.013	–	0.015

Table 3: Decision layer. Win-rate error is the overshoot over a 10% target in bid replay on AuctionNet rows held out within the training periods. Win-probability and surplus errors are on iPinYou at $\lambda = 0.5$, with surplus under the second-price payment rule. Errors other than surplus are in percentage points (pp).

Benchmark	**Metric**	**Naive**	**Censored**	**Corrected**	**Oracle**
AuctionNet replay	win-rate error (pp)	14.3±0.1	17.8±1.6	7.9±2.4	1.8±0.4
iPinYou ($\lambda = 0.5$)	win-probability error (pp)	8.5	6.1±0.7	0.9±0.3	0.4±0.4
iPinYou ($\lambda = 0.5$)	relative surplus error	31%	22±4%	2.6±2.4%	2.2±2.0%

(Figure 2). As the dependence nears one, bids track prices so closely that a win reveals little about the price beyond the bid. Theorem 1 marks this edge, where its positive-density condition fails.

The simplest alternative to the copula feeds the bid to the backbone as one more covariate under the censored loss. That leaves each bid value's price tail to be extrapolated from higher bids, whereas the copula ties every tail to one shared margin. On iPinYou at $\lambda = 0.5$, the covariate fit improves the margin only to 0.035 against the correction's 0.009 (Table 1). DCSurvival, the other alternative, learns the dependence without a parametric family but drops the bid on wins. Its error there is 0.060, no better than the censored network. Once the bid is dropped, fitting the data well no longer identifies the dependence (Table 4, Appendix B.3).

The correction reduces error under the tested misspecifications. The rank mixture of the coupling grid is symmetric in its tails, the shape the Gaussian family suits. Table 2 removes that advantage in two ways. The Clayton and Gumbel rows plant a copula with one heavy tail outside the Gaussian family. The last two rows let the dependence arise from a signal the bidder acts on, so no copula is specified and its strength is measured rather than set. On every row the Gaussian fit removes 74 to 85% of the naive error. On the planted rows, likelihood selection identifies the true family in both datasets and in all 100 replications from either family (Appendix B.4), bringing the margin within 0.001 of the truth. On the signal rows, the fit is within 0.003 of the recursion given the measured dependence on AuctionGym and below it on iPinYou. Across the positive signal shares, the fit removes 73 to 88% of the naive error (Appendix B.4).

The win cells test the family's observable implications. For every bid, the fitted model implies how the won prices should be distributed. Theorem 1 needs only part of the information in these cells, so the rest gives overidentifying restrictions that Pearson's statistic tests against the observed counts. A large value shows a mismatch, but a small one does not confirm the family, because laws outside it can share the same win cells. Fixed-bid records close that gap. Proposition A.3 tests the family against a margin free of the copula (Appendix B.6). Per degree of freedom, Pearson's statistic is near 1.0 under the planted family and several times larger under a wrong one (Appendix B.4). We run it on our fits from the coupling grid. On the iPinYou rank mixture at $\lambda = 0.5$, the estimator without features fails the check in all four fitted families. The check responds to a wrong family and to a shifted dependence alike in the planted replications, without telling which or sizing the CDF error. With the dependence fixed at its true strength, the wrong family leaves little of that error. The proposed model still removes most of the censored network's CDF error there (Table 1).

Corrected distributions reduce decision errors. The learned price distribution decides bid shading, pacing, and the surplus a bidder expects. At a 10% target win rate, corrected bids overshoot the target by 7.9 pp on held-out AuctionNet rows, against 17.8 pp for the censored network and 1.8 for the oracle (Table 3). On iPinYou at $\lambda = 0.5$, the correction cuts the censored network's mean win-probability error from 6.1 to 0.9 pp and its surplus misstatement from 22% to 2.6% (Table 3).

5 Related work

Any method that learns a market-price distribution from its own auction outcomes must account for prices hidden by losses. Bid-landscape models fit censored regressions, survival objectives, or local likelihoods to these outcomes (Wu et al., 2015; Ren et al., 2019; Li et al., 2022). The use of a product-limit estimator on a bidder's logs goes back to Amin et al. (2012). These methods require the price to be independent of the bid after conditioning on the features. No-regret analyses of censored auction feedback state the analogous independence assumption explicitly (Badanidiyuru et al., 2023). Recent auto-bidding work learns a bidding policy, by iterative offline reinforcement learning (Li et al., 2024) or by generating shaded bids that a censored-likelihood price model scores (Huang et al., 2026). Each trains on a log recording the outcome only at the submitted bid, and uses bid exploration for bids that log never tried. Yang et al. (2026) require no exploration and predict how cost and value respond to a bid multiplier over the remaining horizon. The logged multipliers are realized under different campaign histories, which they raise as an identifiability concern for the horizon-aggregate response. They therefore target prediction under the deployment distribution rather than the causal curve. We ask instead what the censored log identifies about the price.

Allowing the bid to depend on the price connects our estimator to copula methods for dependent censoring. The copula-graphic estimator recovers a marginal survival law under a specified copula (Zheng & Klein, 1995). Later work estimates the association itself under restrictions on the margins (Czado & Van Keilegom, 2023), with Schwarz et al. (2013) mapping which parameters a fixed family identifies. DCSurvival learns a neural copula while requiring both margins to satisfy a separate identifiability condition (Zhang et al., 2024). A related scheme appears in survival studies where each record shows a threshold and whether failure preceded it, as the bid does in the auction log. Wang et al. (2012) identify a nonparametric failure law under a fully known copula, whereas Theorem 1 frees the dependence parameter through the exact win-region cells. Appendix C compares Theorem 2 with the estimation guarantees of dependent-censoring and missing-data research.

Without such structure, our boundary is the classical nonidentifiability under dependent censoring and its marginal bounds (Tsiatis, 1975; Peterson, 1976). Missing-data theory reaches the same boundary (Nabi et al., 2020) and lifts it with a shadow variable (Chen et al., 2023) or with selection on observables (Coston et al., 2021). Neither route follows from the auction log. The bid determines censoring itself, while the features need not remove its dependence on the price. Causal sensitivity analysis treats a bound as the answer when an identifying assumption is withheld (Frauen et al., 2024). Proposition 1 gives a sharp band with no sensitivity parameter and an observed width.

6 Discussion

The bidder's own policy decides which prices its log reveals. The censored log alone confines the price margin to a sharp band with a minimax floor. Under affiliation, the band narrows to a closed form and the standard censored fit understates the truth, more so as the dependence strengthens. A concordance-ordered copula family collapses the band to a point through the largest bid. The two-stage estimator carries the rate guarantee. The correction removes three quarters of the censored network's CRPS gap to the oracle, improving the bids, win probabilities, and surplus downstream. Every theorem in the chain rests on one fact of the record, the always-observed bid.

The Gaussian copula remains an identifying assumption. The win-cell check of Section 4 can reject the family but not confirm it, and under a wrong family a better likelihood can come with a worse margin (Appendix B.4). Fixed-bid records give a copula-free margin to check the family against (Proposition A.3). Above the fixed bid, the copula still carries F. The correction also degrades as the dependence nears one, the edge the positive-density condition of Theorem 1 marks. Our diagnostics do not split the remaining quarter among the copula, the conditional dependence, and the network. Validation against production-scale bidding systems is the natural next step.

Ethics statement

This work uses the public AuctionNet and iPinYou benchmarks and a reimplemented AuctionGym simulator, none of which contain personally identifiable information. No human subjects are involved. Our contribution lies only in correcting a statistical bias in market-price estimation. We foresee no harms beyond the existing competitive dynamics of advertising markets and consider the work consistent with the ICLR Code of Ethics.

Reproducibility statement

Complete proofs of every stated result are in Appendix A. All three benchmarks are public. Appendix B.1 states the fitted likelihood, Appendix B.2 records the acquisition, preprocessing, seeds, splits, and training protocol of each benchmark, and Appendix B.3 gives the configuration of every baseline.

AI use statement

We used AI coding agent tools to edit for clarity and to improve notational consistency in the censored-likelihood formalization in Sections 2 and 3. We also used them to provide initial drafts of the proofs in Appendix A.3, each of which we verified line by line against its numerical witness script. We further used them to assist in writing the preprocessing and evaluation pipeline of Appendix B. We also used AI tools to identify relevant literature. We have reviewed all AI-assisted material and take responsibility for the final content of this work.

References

Juliette Achddou, Olivier Cappé, and Aurélien Garivier. Efficient algorithms for stochastic repeated second-price auctions. In *Proceedings of the 32nd International Conference on Algorithmic Learning Theory*, volume 132 of *Proceedings of Machine Learning Research*, pp. 99–150, 2021.

Kareem Amin, Michael Kearns, Peter Key, and Anton Schwaighofer. Budget optimization for sponsored search: Censored learning in MDPs. In *Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*, pp. 54–63. AUAI Press, 2012.

Susan Athey and Philip A. Haile. Identification of standard auction models. *Econometrica*, 70(6): 2107–2140, 2002. doi: 10.1111/j.1468-0262.2002.00435.x.

Ashwinkumar Badanidiyuru, Zhe Feng, and Guru Guruganesh. Learning to bid in contextual first price auctions. In *Proceedings of the ACM Web Conference 2023*, pp. 3489–3497, 2023. doi: 10.1145/3543507.3583427.

George H. Chen. Nearest neighbor and kernel survival analysis: Nonasymptotic error bounds and strong consistency rates. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 1001–1010, 2019.

Jacob M. Chen, Daniel Malinsky, and Rohit Bhattacharya. Causal inference with outcome-dependent missingness and self-censoring. In *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, volume 216 of *Proceedings of Machine Learning Research*, pp. 358–368, 2023.

Amanda Coston, Ashesh Rambachan, and Alexandra Chouldechova. Characterizing fairness over the set of good models under selective labels. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 2144–2155, 2021.

Gilles Crommen, Jad Beyhum, and Ingrid Van Keilegom. An instrumental variable approach under dependent censoring. *TEST*, 33(2):473–495, 2024. doi: 10.1007/s11749-023-00903-9.

Claudia Czado and Ingrid Van Keilegom. Dependent censoring based on parametric copulas. *Biometrika*, 110(3):721–738, 2023. doi: 10.1093/biomet/asac067.

Negera Wakgari Deresa and Ingrid Van Keilegom. On semiparametric modelling, estimation and inference for survival data subject to dependent censoring. *Biometrika*, 108(4):965–979, 2021. doi: 10.1093/biomet/asaa095.

Othman El Balghiti, Adam N. Elmachtoub, Paul Grigas, and Ambuj Tewari. Generalization bounds in the predict-then-optimize framework. In *Advances in Neural Information Processing Systems*, volume 32, pp. 14389–14398. Curran Associates, Inc., 2019.

Ali Hossein Gharari Foomani, Michael Cooper, Russell Greiner, and Rahul G. Krishnan. Copula-based deep survival models for dependent censoring. In *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, volume 216 of *Proceedings of Machine Learning Research*, pp. 669–680, 2023.

Dennis Frauen, Fergus Imrie, Alicia Curth, Valentyn Melnychuk, Stefan Feuerriegel, and Mihaela van der Schaar. A neural framework for generalized causal sensitivity analysis. In *International Conference on Learning Representations*, 2024.

Djordje Gligorijevic, Tian Zhou, Bharatbhushan Shetty, Brendan Kitts, Shengjun Pan, Junwei Pan, and Aaron Flores. Bid shading in the brave new world of first-price auctions. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management*, pp. 2453–2460, 2020. doi: 10.1145/3340531.3412689.

Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.

Anna Guo, Jiwei Zhao, and Razieh Nabi. Sufficient identification conditions and semiparametric estimation under missing not at random mechanisms. In *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, volume 216 of *Proceedings of Machine Learning Research*, pp. 777–787, 2023.

Yinqiu Huang, Hao Ma, Wenshuai Chen, Zongwei Wang, Shuli Wang, Yongqiang Zhang, Xue Wei, Yinhua Zhu, Haitao Wang, and Xingxing Wang. Generative bid shading in real-time bidding advertising. In *Proceedings of the 49th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 600–609, 2026. doi: 10.1145/3805712.3809566.

Olivier Jeunen, Sean Murphy, and Ben Allison. Off-policy learning-to-bid with auctiongym. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 4219–4228, 2023. doi: 10.1145/3580305.3599877.

Edward L. Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481, 1958. doi: 10.1080/01621459.1958.10501452.

Samuel Karlin. *Total Positivity*. Stanford University Press, 1968.

Samuel Karlin and Yosef Rinott. Classes of orderings of measures and related correlation inequalities. I. multivariate totally positive distributions. *Journal of Multivariate Analysis*, 10(4):467–498, 1980. doi: 10.1016/0047-259X(80)90065-2.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015.

Vijay Krishna. *Auction Theory*. Academic Press, second edition, 2010.

Haoming Li, Yusen Huo, Shuai Dou, Zhenzhe Zheng, Zhilin Zhang, Chuan Yu, Jian Xu, and Fan Wu. Trajectory-wise iterative reinforcement learning framework for auto-bidding. In *Proceedings of the ACM Web Conference 2024*, pp. 4193–4203, 2024. doi: 10.1145/3589334.3645534.

Xu Li, Michelle Ma Zhang, Zhenya Wang, and Youjun Tong. Arbitrary distribution modeling with censorship in real-time bidding advertising. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 3250–3258, 2022. doi: 10.1145/3534678.3539048.

Paul R. Milgrom and Robert J. Weber. A theory of auctions and competitive bidding. *Econometrica*, 50(5):1089–1122, 1982.

Razieh Nabi, Rohit Bhattacharya, and Ilya Shpitser. Full law identification in graphical models of missing data: Completeness results. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 7153–7163, 2020.

Whitney K. Newey and Daniel McFadden. Large sample estimation and hypothesis testing. In Robert F. Engle and Daniel L. McFadden (eds.), *Handbook of Econometrics*, volume 4, chapter 36, pp. 2111–2245. North-Holland, Amsterdam, 1994.

Arthur V. Peterson, Jr. Bounds for a joint distribution function with fixed sub-distribution functions: Application to competing risks. *Proceedings of the National Academy of Sciences*, 73(1):11–13, 1976. doi: 10.1073/pnas.73.1.11.

Robin L. Plackett. A reduction formula for normal multivariate integrals. *Biometrika*, 41(3-4): 351–360, 1954. doi: 10.1093/biomet/41.3-4.351.

Kan Ren, Jiarui Qin, Lei Zheng, Zhengyu Yang, Weinan Zhang, and Yong Yu. Deep landscape forecasting for real-time bidding advertising. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 363–372, 2019. doi: 10.1145/3292500.3330870.

Maik Schwarz, Geurt Jongbloed, and Ingrid Van Keilegom. On the identifiability of copulas in bivariate competing risks models. *The Canadian Journal of Statistics*, 41(2):291–303, 2013. doi: 10.1002/cjs.11179.

Moshe Shaked and J. George Shanthikumar. *Stochastic Orders*. Springer, 2007.

Abe Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8:229–231, 1959.

Eric V. Slud and L. V. Rubinstein. Dependent competing risks and summary survival curves. *Biometrika*, 70(3):643–649, 1983. doi: 10.2307/2336500.

Kefan Su, Yusen Huo, Zhilin Zhang, Shuai Dou, Chuan Yu, Jian Xu, Zongqing Lu, and Bo Zheng. Auctionnet: A novel benchmark for decision-making in large-scale games. In *Advances in Neural Information Processing Systems*, volume 37, 2024. Datasets and Benchmarks Track.

Herbert Thijs, Geert Molenberghs, Bart Michiels, Geert Verbeke, and Desmond Curran. Strategies to fit pattern-mixture models. *Biostatistics*, 3(2):245–265, 2002. doi: 10.1093/biostatistics/3.2.245.

Anastasios Tsiatis. A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, 72(1):20–22, 1975. doi: 10.1073/pnas.72.1.20.

Chunjie Wang, Jianguo Sun, Liuquan Sun, Jie Zhou, and Dehui Wang. Nonparametric estimation of current status data with dependent censoring. *Lifetime Data Analysis*, 18(4):434–445, 2012. doi: 10.1007/s10985-012-9223-7.

Wush Chi-Hsuan Wu, Mi-Yen Yeh, and Ming-Syan Chen. Predicting winning price in real time bidding with censored data. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1305–1314, 2015. doi: 10.1145/2783258.2783276.

Jian Xu, Kuang-chih Lee, Wentong Li, Hang Qi, and Quan Lu. Smart pacing for effective online ad campaign optimization. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 2217–2226, 2015. doi: 10.1145/2783258.2788615.

Eunseok Yang, Xingdong Zuo, and Kyung-Min Kim. Constrained auto-bidding via generative response modeling. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*, pp. 5975–5985, 2026. doi: 10.1145/3770855.3817847.

Weijia Zhang, Chun Kai Ling, and Xuanhui Zhang. Deep copula-based survival analysis for dependent censoring with identifiability guarantees. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 20613–20621, 2024. doi: 10.1609/aaai.v38i18.30047.

Weinan Zhang, Shuai Yuan, and Jun Wang. Optimal real-time bidding for display advertising. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1077–1086, 2014a. doi: 10.1145/2623330.2623633.

Weinan Zhang, Shuai Yuan, Jun Wang, and Xuehua Shen. Real-time bidding benchmarking with ipinyou dataset, 2014b. arXiv:1407.7073; UCL Technical Report.

Ming Zheng and John P. Klein. Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika*, 82(1):127–138, 1995. doi: 10.1093/biomet/82.1.127.

Tian Zhou, Hao He, Shengjun Pan, Niklas Karlsson, Bharatbhushan Shetty, Brendan Kitts, Djordje Gligorijevic, San Gultekin, Tingyu Mao, Junwei Pan, Jianlong Zhang, and Aaron Flores. An efficient deep distribution network for bid shading in first-price auctions. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 3996–4004, 2021. doi: 10.1145/3447548.3467167.

A PROOFS

This appendix proves the results of Section 2 in the order they appear. It also states and proves the two supporting propositions referenced there.

A.1 INFORMATION BOUNDARY

Proof of Proposition 1. The observable law consists of the win-region cells $\pi(z, b)$ with $z \leq b$ and the loss masses $\ell(b)$ of Section 2. These satisfy

$$\ell(b) = \mathbb{P}(B{=}b) - \sum_{z \leq b} \pi(z, b),$$

so the bid margin and the win-region cells determine the observable law. Conversely, the observable law contains each of these probabilities. This proves the equivalence claim. The likelihood of any dataset is then shared by two such laws, since it is a product of win-region cells $\pi(z_i, b_i)$ and loss masses $\ell(b_i)$.

Fix $k < K$. Partitioning on W, we have

$$F(k) = \mathbb{P}(W{=}1, Z \leq k) + \mathbb{P}(W{=}0, Z \leq k).$$

The first term is observed. Since a loss implies $Z > B$, a loss with $B \geq k$ must have $Z > k$ and cannot contribute to $F(k)$. Only losses with $B < k$ remain ambiguous. Therefore,

$$\mathbb{P}(W{=}1, Z \leq k) \leq F(k) \leq \mathbb{P}(W{=}1, Z \leq k) + \mathbb{P}(W{=}0, B < k).$$

Place every lost price at K to attain the lower bound. For the upper bound, place the loss mass at bid b at $b + 1$. Both completions satisfy $Z > B$ and preserve the observable law. Neither depends on k, so the two envelopes are attained simultaneously for all $k < K$. □

Proof of Corollary 1. Because P and P' induce the same observable law, the estimator $\hat{F}$ has the same distribution under both. For every realization the triangle inequality gives

$$\Delta = \|F - F'\|_\infty \leq \|\hat{F} - F\|_\infty + \|\hat{F} - F'\|_\infty.$$

Taking the expectation under P, and moving the second term to P' through the shared distribution of $\hat{F}$,

$$\Delta \leq \mathbb{E}_P\|\hat{F} - F\|_\infty + \mathbb{E}_{P'}\|\hat{F} - F'\|_\infty \leq 2 \max_{Q \in \{P, P'\}} \mathbb{E}_Q\|\hat{F} - F_Q\|_\infty.$$

□

A.2 NAIVE BIAS

Proof of Proposition 2. Fix k and write $w_b = \mathbb{P}(Z \geq k, B{=}b)$. Bayes' rule gives the two hazards as weighted averages of the same conditional hazards $q_b(k)$,

$$h(k) = \frac{\sum_b w_b\, q_b(k)}{\sum_b w_b}, \qquad h^{\#}(k) = \frac{\sum_{b\geq k} w_b\, q_b(k)}{\sum_{b\geq k} w_b}.$$

If $q_b(k)$ is nonincreasing in b, every excluded term ($b < k$) satisfies $q_b(k) \geq q_{b'}(k)$ for every included $b' \geq k$, so the restricted average cannot exceed the full one: $h^{\#}(k) \leq h(k)$. Since $0 \leq h^{\#}(j) \leq h(j)$ for all $j \leq k$, the survival products satisfy $\prod_{j\leq k}(1-h^{\#}(j)) \geq \prod_{j\leq k}(1-h(j))$, i.e. $F^{\#}(k) \leq F(k)$.

Under affiliation, the joint pmf is totally positive. It follows that $Z \mid B{=}b$ increases in the likelihood-ratio order as b increases. Likelihood-ratio order implies hazard-rate order, which is the required pointwise ordering of $q_b(k)$ (Shaked & Shanthikumar, 2007). □

Proposition A.1 (Exact bias identity)**.** *For every k with $\mathbb{P}(Z \geq k,\, B \geq k) > 0$ and $\mathbb{P}(Z \geq k,\, B < k) > 0$,*

$$h(k) - h^{\#}(k) \;=\; \alpha(k)\,\big(A(k) - h^{\#}(k)\big),$$

where $\alpha(k) = \mathbb{P}(B < k \mid Z \geq k)$ and $A(k) = \mathbb{P}(Z{=}k \mid Z \geq k,\, B < k)$. Consequently $|F(k) - F^{\#}(k)| \leq \sum_{j\leq k} |h(j) - h^{\#}(j)|$ for every k.

Proof. Partition the events $\{Z = k\}$ and $\{Z \geq k\}$ by whether $B \geq k$. Their conditional mixture is $h(k) = (1-\alpha(k))\, h^{\#}(k) + \alpha(k)\, A(k)$. Rearranging proves the identity. For the CDF bound, with $a_j = 1 - h^{\#}(j)$ and $b_j = 1 - h(j)$ in $[0,1]$,

$$\textstyle\prod_{j=1}^{k} a_j - \prod_{j=1}^{k} b_j = \sum_{j=1}^{k} \big(\prod_{i=1}^{j-1} a_i\big)(a_j - b_j)\big(\prod_{i=j+1}^{k} b_i\big),$$

and every prefactor lies in $[0,1]$, so $|F(k) - F^{\#}(k)| \leq \sum_{j\leq k} |h(j) - h^{\#}(j)|$. □

Proposition A.2 (Bias monotone in the dependence)**.** *Couple (Z, B) with fixed margins through a family $\{C_\theta : \theta \in \Theta\}$ that is differentiable in θ on the interior of Θ and strictly ordered by concordance. Let θ_{ind} be its lower endpoint, and suppose $C_{\theta_{\mathrm{ind}}}(u,v) = uv$. Write $\bar{S}_\theta(u,v) = 1 - u - v + C_\theta(u,v)$. For any k at which the two survival terms below are positive,*

$$1 - h^{\#}_\theta(k) \;=\; \frac{\bar{S}_\theta\big(F(k),\, G(k{-}1)\big)}{\bar{S}_\theta\big(F(k{-}1),\, G(k{-}1)\big)},$$

and on any parameter interval where

$$\partial_\theta \log \bar{S}_\theta\big(F(k), G(k{-}1)\big) \;\geq\; \partial_\theta \log \bar{S}_\theta\big(F(k{-}1), G(k{-}1)\big) \quad \textit{(condition } (L_k)\textit{)}$$

the crude hazard $h^{\#}_\theta(k)$ is nonincreasing in θ. If, for every $j \leq k$, both corresponding survival terms are positive and (L_j) holds, the population product-limit $F^{\#}_\theta(k)$ is nonincreasing in θ, with $F^{\#}_{\theta_{\mathrm{ind}}}(k) = F(k)$ wherever the crude hazards are defined.

Proof. For $j \in \{k, k+1\}$, the upper-orthant probability is

$$\mathbb{P}(Z \geq j,\, B \geq k) = \bar{S}_\theta(F(j{-}1), G(k{-}1)).$$

Taking their ratio gives

$$1 - h^{\#}_\theta(k) = \mathbb{P}(Z \geq k{+}1 \mid Z \geq k,\, B \geq k) = \frac{\bar{S}_\theta(F(k), G(k{-}1))}{\bar{S}_\theta(F(k{-}1), G(k{-}1))}.$$

Under (L_k),

$$\partial_\theta \log\big(1 - h^{\#}_\theta(k)\big) = \partial_\theta \log \bar{S}_\theta(F(k), G(k{-}1)) - \partial_\theta \log \bar{S}_\theta(F(k{-}1), G(k{-}1)) \;\geq\; 0,$$

so each product-limit factor $1 - h^{\#}_\theta(j)$, $j \leq k$, is nondecreasing in θ and $F^{\#}_\theta(k) = 1 - \prod_{j\leq k}(1 - h^{\#}_\theta(j))$ is nonincreasing. At independence, $C_{\theta_{\mathrm{ind}}}(u,v) = uv$ and $\bar{S}_{\theta_{\mathrm{ind}}}(u,v) = (1-u)(1-v)$. Hence

$$1 - h^{\#}_{\theta_{\mathrm{ind}}}(k) = \frac{1 - F(k)}{1 - F(k-1)}.$$

The product telescopes to $1 - F(k)$, proving $F^{\#}_{\theta_{\text{ind}}}(k) = F(k)$. □

Condition (L_k) can be checked directly. On the benchmark margins, the Gaussian family satisfies it at every bucket and parameter step except the last, where Kendall's τ reaches 0.7 or higher. Those violations occur in the first buckets above the minimum bid and are at most 4.5×10^{-4}, against drops of 0.19–0.29 in the population product-limit along the parameter path.

A.3 IDENTIFICATION

The bounds of Theorem 1(i). For a support point $b < b_{\max}$, write b^+ for the next support point and $\kappa_b = \min\{\pi(z,b)/\pi(z,b^+) : z \le b,\ \pi(z,b^+) > 0\}$. For a price $j \le b_{\max}$, let $d(j) = q_b(j)$ be the conditional hazard of Proposition 2 at the smallest support point $b \ge j$, a ratio of observed cells. With sums over the support points $b < k$,

$$F_L(k) = \mathbb{P}(W{=}1, Z \le k) + \sum_{b<k} \ell(b)\Big[1 - \prod_{j=b+1}^{k} \big(1 - d(j)\big)\Big],$$
$$F_U(k) = \mathbb{P}(W{=}1, Z \le k) + \sum_{b<k} M_b(k),$$

where

$$M_b(k) = \min\Big\{\ell(b),\ \kappa_b\big[\mathbb{P}(b < Z \le \min(k, b^+),\ B{=}b^+) + \mathbf{1}\{b^+ < k\}\, M_{b^+}(k)\big]\Big\}.$$

Proof of Theorem 1(i). Write $c = b^+$. A completion assigns to each lost bid b a tail $t_b(z) \ge 0$, $z > b$, summing to $\ell(b)$. Every completion preserves the observable law, and since a loss at b has $Z > b$,

$$F(k) = \mathbb{P}(W{=}1,\ Z \le k) + \sum_{b<k} T_b(k), \qquad T_b(k) = \sum_{z=b+1}^{k} t_b(z). \tag{1}$$

Affiliation places column b below column c in the likelihood-ratio order, $\pi(z,b)\pi(z',c) \ge \pi(z',b)\pi(z,c)$ for $z < z'$, and the order is transitive, so consecutive support points suffice. Rows at or below b are observed and already ordered. On rows $z \le b < z'$, this says $t_b(z') \le \kappa_b\, \pi(z',c)$, and on rows $b < z < z'$, it says that $t_b/\pi(\cdot,c)$ is nonincreasing. Summing the first over $z' > b$ gives $\ell(b) \le \kappa_b\, \mathbb{P}(Z > b,\ B{=}c)$.

Lower edge. The likelihood-ratio order implies the hazard-rate order (Shaked & Shanthikumar, 2007), so for a support point $b < k$ and every price $j \in (b,k]$, the hazard of column b is at least that of the smallest support point at or above j, which is $d(j)$. A loss at b with $Z > k$ survives every price in $(b,k]$, so $\mathbb{P}(Z > k, B{=}b) \le \ell(b) \prod_{j=b+1}^{k}(1 - d(j))$, and equation 1 gives $F(k) \ge F_L(k)$. The bound is attained by continuing every lost auction with the hazards $d(j)$. Put $\ell(b_{\max})$ in row K and, going down the support points, make t_b proportional to column c on $(b,K]$. The ratio $\ell(b)/\mathbb{P}(Z > b, B{=}c)$ is at most κ_b by the inequality above, and consecutive columns share their conditional law above b, so the completion is affiliated and each lost column carries the hazards $d(j)$ above b. This is the neighboring-case completion of pattern-mixture models (Thijs et al., 2002), and the bound is that of Slud & Rubinstein (1983) with their hazard ratio at the value affiliation allows.

Upper edge. Summing $t_b(z') \le \kappa_b \pi(z',c)$ over $z' \in (b,k]$ gives

$$T_b(k) \le \kappa_b\big[\mathbb{P}(b < Z \le \min(k,c), B{=}c) + \mathbf{1}\{c < k\}\, T_c(k)\big],$$

since the rows of column c up to c are observed and the rest is its tail. With $T_b(k) \le \ell(b)$, induction downward from the largest support point below k gives $T_b(k) \le M_b(k)$, and equation 1 gives $F(k) \le F_U(k)$. The bound is attained by front-loading. Put $\ell(b_{\max})$ in row $b_{\max} + 1$ and, going down the support points, fill each column from row $b+1$ upward with $t_b(z) = \min\{\kappa_b \pi(z,c),\ \text{remaining mass}\}$, which places all of $\ell(b)$ by the inequality above. The ratio $t_b/\pi(\cdot,c)$ is κ_b, at most κ_b in the row where the mass runs out, and zero after, so the completion is affiliated, and its partial sums are $\min\{\ell(b), \kappa_b \sum_{z=b+1}^{k} \pi(z,c)\}$, which is the recursion for $M_b(k)$. Neither completion depends on k. □

Remark A.1 (Family members inside the band). If the copula density of a member of the family in part (ii) is totally positive of order two, as the Gaussian copula density is for $\rho \geq 0$ (Karlin & Rinott, 1980), then the law it induces on the grid is affiliated, and the point of part (ii) lies in the band of part (i).

Proof. Each cell is $\pi(z, b) = \iint c_\theta(u, v)\, \mathbf{1}\{u \in I_z\}\mathbf{1}\{v \in J_b\}\, du\, dv$ with $I_z = (F(z{-}1), F(z)]$ and $J_b = (G(b{-}1), G(b)]$. The indicator of ordered intervals is totally positive in (u, z), and likewise in (v, b), so two applications of the basic composition formula (Karlin, 1968) make π totally positive of order two in (z, b). □

Proof of Theorem 1(ii). The always-observed bid identifies G. A win reveals the price, so the log also identifies the probability of winning at each price

$$p_k = \mathbb{P}(Z{=}k, B \geq k),$$

and the probability of winning with the smallest bid

$$a_0 = \mathbb{P}(Z \leq b_0, B{=}b_0),$$

observable because $Z \leq b_0 = B$ implies $W = 1$. The recovery runs in three steps, first F up to the smallest bid, then θ from one equation, then F one price at a time.

Step 1. For $k \leq b_0$, every bid satisfies $B \geq k$, so no bid can censor these prices and $p_k = \mathbb{P}(Z{=}k)$. Hence

$$F(k) = \sum_{j=1}^{k} p_j, \qquad k \leq b_0.$$

In particular, the log identifies $F(b_0)$.

Step 2. Since b_0 is the smallest bid support point, the events $\{B \leq b_0\}$ and $\{B{=}b_0\}$ coincide, and

$$a_0 = \mathbb{P}(Z \leq b_0, B \leq b_0) = C_\theta\big(F(b_0), G(b_0)\big).$$

Every quantity in this equation except θ is now known. Since $0 < F(b_0) < 1$ and $b_0 < b_{\max}$ place both copula arguments inside $(0, 1)$, strict ordering in θ makes the equation one-to-one, and θ is determined.

Step 3. Fix $b_0 < k \leq b_{\max}$ and define

$$r_k(u) = u - C_\theta\big(u, G(k-1)\big).$$

Splitting the event $\{Z{=}k\}$ by whether the bid reaches k, the probability of winning at price k obeys

$$\begin{aligned} p_k &= \mathbb{P}(Z{=}k) - \mathbb{P}(Z{=}k, B \leq k{-}1) \\ &= F(k) - F(k-1) \\ &\quad - [C_\theta(F(k), G(k-1)) - C_\theta(F(k-1), G(k-1))] \\ &= r_k(F(k)) - r_k(F(k-1)). \end{aligned}$$

The only unknown here is $F(k)$. For $0 \leq u_1 < u_2 \leq 1$ the definition of r_k expands to

$$r_k(u_2) - r_k(u_1) = (u_2 - u_1) - \big[C_\theta(u_2, G(k-1)) - C_\theta(u_1, G(k-1))\big],$$

and absolute continuity writes both pieces as integrals of the density, with $v = G(k-1)$,

$$u_2 - u_1 = \int_{u_1}^{u_2}\int_0^1 c_\theta(s, t)\, dt\, ds, \qquad C_\theta(u_2, v) - C_\theta(u_1, v) = \int_{u_1}^{u_2}\int_0^{v} c_\theta(s, t)\, dt\, ds.$$

Subtracting the second from the first leaves the rectangle $[u_1, u_2] \times [G(k-1), 1]$, and $k \leq b_{\max}$ keeps $G(k-1) < 1$, so the strictly positive density gives

$$r_k(u_2) - r_k(u_1) = \int_{u_1}^{u_2}\int_{G(k-1)}^1 c_\theta(s, t)\, dt\, ds > 0.$$

Hence r_k is strictly increasing, and

$$F(k) = r_k^{-1}(p_k + r_k(F(k-1)))$$

is uniquely determined. Starting from $F(b_0)$ and repeating this step recovers the price margin through $b_{\max}$.

For $k > b_{\max}$, we have $G(k-1) = 1$ and $r_k(u) = u - C_\theta(u, 1) = 0$. The recursion then imposes no restriction on $F(k)$, matching the boundary in Proposition 1. □

Proof of Corollary 2. Suppose first that two laws of the family, with margins (F, G) and (F', G') and parameters θ and θ', are observationally equivalent. The always-observed bid gives $G = G'$. The win-region cells also agree. The recursion in the proof of Theorem 1 therefore returns the same output for both laws, giving $\theta = \theta'$ and $F(k) = F'(k)$ for all $k \le b_{\max}$. Conversely, every observable cell is a function of θ, G, and $F(1), \dots, F(b_{\max})$ alone. A win cell with $z \le b \le b_{\max}$ is the copula rectangle built from $F(z)$, $F(z{-}1)$, $G(b)$, $G(b{-}1)$. The loss cell at bid b is $g(b)$ minus the win cells in its column. Two laws of the family sharing these inputs are therefore observationally equivalent.

At $k \le b_{\max}$ the identified set within the family is the range of $F(k)$ over laws observationally equivalent to the truth, a single point by the first part. Every pair used in the proof of Corollary 1 must be observationally equivalent, so within the family such a pair satisfies $F = F'$ on $\{1, \dots, b_{\max}\}$ and its restricted sup distance is zero. At $b_{\max} < k < K$, replacing F beyond $b_{\max}$ by any nondecreasing extension with $F'(K) = 1$ yields a valid law of the family with the same observable cells, so every value in $[F(b_{\max}), 1]$ is attained, while monotonicity forces $F(k) \ge F(b_{\max})$. □

Remark (recovery without the smallest-bid anchor). Step 2 uses $0 < F(b_0) < 1$. If $F(b_0) = 0$, let $z^* \le b_{\max}$ be the first price with $F(z^*) > 0$ and b_1 the smallest bid atom at or above it. The win cells $S_1 = \mathbb{P}(Z = z^*, B \ge z^*)$ and $S_2 = \mathbb{P}(Z = z^*, B = b_1)$ satisfy

$$S_1 = u - C_\theta(u, v_1), \qquad S_2 = C_\theta(u, v_2) - C_\theta(u, v_1),$$

with $u = F(z^*)$, $v_1 = G(z^*{-}1) < 1$ and $v_2 = G(b_1)$. Under positive density the first equation fixes $u(\theta)$; the second recovers θ whenever the profile $\theta \mapsto C_\theta(u(\theta), v_2) - C_\theta(u(\theta), v_1)$ is injective, after which Step 3 runs from z^*. The profile is void when $b_1 = b_{\max}$, since then $v_2 = 1$ and $S_2 = S_1$. We check the profile numerically on the iPinYou bid margin for the four families of Section 4, and it is monotone in every configuration with a nonempty win cell.

Fixed-bid records. A bidder can hold a random share of auctions at a fixed bid t and mark them in the log, lost as well as won. The mark is the indicator R below, independent of the price given the targeting features because the share is drawn at random.

Proposition A.3 (Fixed-bid records)**.** *Let $R \in \{0, 1\}$ mark a share of auctions in the log, with $0 < \mathbb{P}(R{=}1) < 1$, drawn without regard to the price or to the inputs of the bidding rule. On marked auctions, the bid equals a fixed t. On the rest, (Z, B) is coupled through a family satisfying the copula conditions of Theorem 1(ii), with bid margin G_0 and largest support point $b_{\max}$. Both share the price margin F. If $0 < F(t) < 1$ and $0 < G_0(t-1) < 1$, the log determines θ and $F(k)$ for every $k \le b_{\max}$, with the condition $0 < F(b_0) < 1$ of Theorem 1(ii) replaced by these conditions at t. Moreover, every $s \le t$ with $0 < F(s) < 1$ and $0 < G_0(s-1) < 1$ gives its own equation for θ, whose roots coincide under the family.*

The marked auctions alone give F up to t. Above t, F follows from the recursion of Theorem 1(ii) on the unmarked auctions (Step 3 below), so the copula remains the identifying assumption there. The size of the share affects precision, not identification, so the share can be small. Checking the family needs at least two such s.

Proof of Proposition A.3. The proof follows that of Theorem 1(ii), with the fixed-bid records in place of the prices at or below b_0.

Step 1. On $R{=}1$, every bid equals t, so a price $Z \le t$ is always won and recorded. Because R is drawn without regard to the price and F is the same under both regimes,

$$F(k) = \mathbb{P}(W{=}1,\, Z \le k \mid R{=}1), \qquad k \le t.$$

Step 2. On $R{=}0$, the always-observed bid identifies G_0. An auction with $Z \le t \le B$ is won, so

$$S_t = \mathbb{P}(Z \le t,\, B \ge t \mid R{=}0)$$

is a sum of win cells. Removing from $\{Z \le t\}$ the auctions whose bid falls below t gives

$$S_t = F(t) - C_\theta\big(F(t),\, G_0(t-1)\big).$$

Both copula arguments lie in $(0, 1)$. Strict ordering in θ then makes the right side strictly decreasing in θ, so S_t determines θ.

Step 3. Since $G_0(t-1) < 1$, we have $t \le b_{\max}$. Starting from $F(t)$, Step 3 of the proof of Theorem 1(ii) runs unchanged on the win cells of $R{=}0$ for $k = t+1, \dots, b_{\max}$. No step uses the cell at the smallest bid, so $F(b_0) = 0$ is allowed.

Equality restrictions. Step 1 gives $F(s)$ for every $s \le t$. At each such s with $0 < F(s) < 1$ and $0 < G_0(s-1) < 1$, Step 2 gives the equation $S_s = F(s) - C_\theta(F(s), G_0(s-1))$, whose root is θ under the family. A population whose equations have different roots, or no root, therefore lies outside the family. Equal roots do not confirm the family, since the equations involve the copula at the points $(F(s), G_0(s-1))$ with $s \le t$ and nowhere else. □

A.4 Estimator guarantee

Regularity conditions for Theorem 2.

- (R1) The auctions are i.i.d. from a member of a copula family satisfying the hypotheses of Theorem 1(ii), the bid support $\mathcal{B} = \{b_1 < \cdots < b_q = b_{\max}\}$ has $g_0(b_r) > 0$ for all r, and $b_{\max} < K$.
- (R2) θ_0 is interior to a compact Θ, and the true hazards lie in $(\delta, 1-\delta)$.
- (R3) The observable-cell probabilities are continuous and bounded away from zero on $\Theta \times \mathcal{F}_\delta$ for bid margins in a compact neighborhood of g_0, and twice continuously differentiable in the stage-one and stage-two parameters near the truth.
- (R4) The negative expected Hessian $\mathcal{I}$ of the stage-two log-likelihood is nonsingular at the truth.

Proof of Theorem 2. Let

$$O_i = (B_i, W_i, W_i Z_i)$$

denote the observed auction. For any observed auction O, let $c(O)$ be its cell, namely (Z, B) after a win and the loss cell indexed by B otherwise. Write

$$\gamma = \big(g(b_1), \dots, g(b_{q-1})\big), \qquad g(b_q) = 1 - \sum_{r<q} g(b_r),$$

so that γ gives free coordinates for the bid margin. The first-stage estimate $\hat\gamma$ consists of the corresponding empirical bid frequencies, and γ_0 denotes the true value.

Let $p_c(\eta, \gamma)$ be the probability of observable cell c, including the bid CDF induced by γ. If $\hat p_c$ is its empirical frequency, the normalized stage-two criterion is

$$Q_n(\eta, \hat\gamma) = \sum_c \hat p_c \log p_c(\eta, \hat\gamma).$$

Consistency. Choose a compact neighborhood $\mathcal{N}$ of γ_0 on which all bid atoms remain positive. With probability tending to one, $\hat\gamma$ belongs to $\mathcal{N}$. On $\Theta \times \mathcal{F}_\delta \times \mathcal{N}$, each modeled observable cell has positive probability. Continuity and compactness then give a common positive lower bound for these probabilities.

The cell space is finite. The empirical cell frequencies converge to their population values, while $\hat\gamma \to \gamma_0$. It follows that

$$\sup_{\eta \in \Theta \times \mathcal{F}_\delta} |Q_n(\eta, \hat\gamma) - Q(\eta)| \xrightarrow{p} 0,$$

where

$$Q(\eta) = \sum_c p_c^0 \log p_c(\eta, \gamma_0), \qquad p_c^0 = p_c(\eta_0, \gamma_0).$$

Let $p^0 = (p_c^0)_c$ and $p(\eta, \gamma_0) = (p_c(\eta, \gamma_0))_c$. Gibbs' inequality gives

$$Q(\eta) - Q(\eta_0) = -D_{\mathrm{KL}}\big(p^0 \,\big\|\, p(\eta, \gamma_0)\big) \le 0.$$

Equality requires all observable cell probabilities to match, in particular the win-region cells at the true bid margin. Theorem 1(ii) then gives $\eta = \eta_0$. The argmax theorem yields $\hat\eta \to \eta_0$ in probability (Newey & McFadden, 1994).

Asymptotic normality. Define the stage-two score

$$s(O;\eta,\gamma) = \partial_\eta \log p_{c(O)}(\eta,\gamma)$$

and let

$$\mathcal{I} = -\mathbb{E}[\partial_\eta s(O;\eta_0,\gamma_0)]\,, \qquad \Gamma = \mathbb{E}[\partial_\gamma s(O;\eta_0,\gamma_0)]\,.$$

The empirical bid margin has influence function

$$\psi_\gamma(O) = (\mathbf{1}\{B = b_r\} - g_0(b_r))_{r=1}^{q-1}\,,$$

so

$$\sqrt{n}(\hat\gamma - \gamma_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_\gamma(O_i).$$

Because η_0 is interior, consistency makes $\hat\eta$ an interior maximizer with probability tending to one. Expanding its score equation at (η_0,γ_0) gives

$$\begin{aligned} 0 = \frac{1}{n}\sum_{i=1}^{n} s(O_i;\eta_0,\gamma_0) - \mathcal{I}(\hat\eta - \eta_0) \\ + \Gamma(\hat\gamma - \gamma_0) + o_p(\|\hat\eta - \eta_0\| + \|\hat\gamma - \gamma_0\|)\,. \end{aligned}$$

The empirical score and $\hat\gamma - \gamma_0$ are $O_p(n^{-1/2})$. Nonsingularity of $\mathcal{I}$ then gives $\hat\eta - \eta_0 = O_p(n^{-1/2})$. Substitution yields

$$\sqrt{n}(\hat\eta - \eta_0) = \mathcal{I}^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}[s(O_i;\eta_0,\gamma_0) + \Gamma\psi_\gamma(O_i)] + o_p(1).$$

The summands are bounded i.i.d. vectors. The multivariate central limit theorem applies with

$$\Omega = \mathrm{Var}(s(O;\eta_0,\gamma_0) + \Gamma\psi_\gamma(O))\,.$$

Because the cell space is finite, $\mathcal{I}$, Γ, and Ω are finite sums. Replacing them by their sample analogues gives the sandwich covariance used in the coverage experiment. □

A.5 Conditional model

The proposed model of Section 3 couples conditional margins $F(\cdot \mid x)$ and $G(\cdot \mid x)$ through one dependence parameter θ common to every x. Two statements carry the results of Section 2 to it. Write $b_0(x)$ and $b_{\max}(x)$ for the smallest and largest support points of $G(\cdot \mid x)$.

Proposition A.4 (Identification with features)**.** *Fix a copula family satisfying the hypotheses of Theorem 1 and couple $Z \sim F(\cdot \mid x)$ and $B \sim G(\cdot \mid x)$ through C_θ with θ common to all x. Suppose some $x^\star$ has $b_0(x^\star) < b_{\max}(x^\star)$ and $0 < F(b_0(x^\star) \mid x^\star) < 1$. Then the observable law determines θ. At every x it then determines $F(k \mid x)$ for all $k \le b_{\max}(x)$.*

Proof. Given $X = x$, the observable law is that of Theorem 1 with margins $F(\cdot \mid x)$ and $G(\cdot \mid x)$. Steps 1 and 2 of that proof at $x^\star$ recover $F(k \mid x^\star)$ for $k \le b_0(x^\star)$ and then θ, the two hypotheses on $x^\star$ being exactly what Step 2 needs. With θ known, Steps 1 and 3 at any x recover $F(k \mid x)$ for $k \le b_{\max}(x)$ without any condition on x. □

Proposition A.5 (Estimator guarantee with finitely many feature values)**.** *Let X take values in a finite set $\mathcal{X}$ with $\mathbb{P}(X{=}x) > 0$ for each x. Write*

$$\eta = \big(\theta,\ \{F(k \mid x)\}_{x\in\mathcal{X},\ k\le b_{\max}(x)}\big).$$

Stage one estimates each $g(\cdot \mid x)$ by the empirical bid frequencies at x. Stage two maximizes the conditional observed-data likelihood of Section 3 over η, with the hazards of every $F(\cdot \mid x)$ constrained to $[\delta, 1-\delta]$ on $\{1,\dots,b_{\max}(x)\}$. Suppose (R1)–(R4) hold for the conditional cell law with this η, with Proposition A.4 supplying the identification in (R1). Then $\hat\eta \to \eta_0$ in probability and $\sqrt{n}(\hat\eta - \eta_0) \to \mathcal{N}(0, \mathcal{I}^{-1}\Omega\mathcal{I}^{-1})$ in distribution, with $\mathcal{I}$ and Ω defined as in Theorem 2 for the conditional cells. If a network class can represent every table $\{G(b \mid x)\}$ and $\{F(k \mid x)\}$ on $\mathcal{X}$, the proposed fit of Section 3 is this maximizer.

Proof. With $\mathcal{X}$ finite the observable cells (x, c) are finitely many and η is finite dimensional, so the proof of Theorem 2 applies with γ collecting the free bid frequencies at every x and ψ_γ the stacked per-x multinomial influence functions. In the uniqueness step, equality in Gibbs' inequality matches every cell probability at every x, and Proposition A.4 turns that into $\eta = \eta_0$. The last claim holds because each stage maximizes its objective over a class containing every table on $\mathcal{X}$, so the network maximizer coincides with the saturated one. □

B Implementation and experimental details

B.1 Stage-two likelihood

Let n_{zb} count wins at price z and bid b, let n_b^{loss} count losses at bid b, and set $n_b = \sum_{z \le b} n_{zb} + n_b^{\text{loss}}$. Stage one gives the empirical pmf $\hat{g}(b) = n_b/n$ and CDF $\hat{G}$ of Section 3. For a row with bid b, write $F_{c,\hat{G}}(k \mid b)$ for the exact conditional price CDF displayed there with x dropped. Conditioning on the discrete event $B = b$ restricts the copula to the bid cell $(\hat{G}(b-1), \hat{G}(b)]$. Hence the observable joint probabilities are

$$\begin{aligned}
\pi_{\theta,F,\hat{G}}(z,b) &:= \Pr_{\theta,F,\hat{G}}(Z = z, B = b) \\
&= \hat{g}(b)\Big[F_{c,\hat{G}}(z \mid b) - F_{c,\hat{G}}(z-1 \mid b)\Big], \\
\ell_{\theta,F,\hat{G}}(b) &:= \Pr_{\theta,F,\hat{G}}(Z > b, B = b) \\
&= \hat{g}(b)\Big[1 - F_{c,\hat{G}}(b \mid b)\Big].
\end{aligned}$$

A win is an exact copula rectangle. A loss is the remaining mass in its bid cell.

Stage two of Section 3 maximizes the conditional objective, which is L_n with the factor $\hat{g}(b)$ dropped from every log term. Since $\hat{g}$ is fixed after stage one, the two objectives differ only by the constant $\sum_b n_b \log \hat{g}(b)$ and hence share the same maximizer. Theorem 2 covers L_n.

For the Gaussian copula, both the estimator without features and the proposed model compute the exact bid-cell conditional

$$\widetilde{F}_{c,\hat{G}}(k \mid b) = \frac{1}{\hat{g}(b)} \int_{\hat{G}(b-1)}^{\hat{G}(b)} \Phi\left(\frac{\Phi^{-1}(F(k)) - \theta\,\Phi^{-1}(u)}{\sqrt{1-\theta^2}}\right) du,$$

the copula rectangle of the cell divided by its mass. Here Φ is the standard normal CDF.

B.2 Benchmarks and training protocol

Shared training protocol. On each benchmark the censored, oracle, and corrected networks share one backbone and differ only in the likelihood. Embeddings have dimension 16, the hidden widths are 256–256 and the batch size is 4096. Adam trains the network weights at rate 10^{-3} and the dependence scalar at 0.03, with the output bias starting at hazard 0.05 and the dependence at independence, $t = 0$. Training runs three epochs on the fixed AuctionNet benchmark and whole epochs up to 1500 updates elsewhere. Neural results use ten torch seeds, 2027–2036, each seeding its own held-out split. Shares of a gap closed are computed within each seed and then averaged. The iPinYou and AuctionGym datasets score a tie as a loss, so a lost bid b contributes the mass at or above b and the loss term of Section 3 is $1 - F_{c,\hat{G}}(b-1 \mid x, b)$ on those two benchmarks. The copula fit clamps the CDF to $[10^{-6}, 1-10^{-6}]$ and each cell mass at 10^{-9} before the logarithm, whereas the censored network runs in log-hazard space without a clamp. This is why the fit with the dependence fixed at zero sits 0.003 below the censored network on AuctionNet in Table 4, although the two objectives coincide in exact arithmetic.

Fit without features. The estimator without features replaces the network by one free hazard logit a_j per price bucket, initialized at $-3 + \epsilon_j$ with $\epsilon_j \sim \mathcal{N}(0, 0.01^2)$, and shares the dependence parameterization and its start. Full-batch Adam (Kingma & Ba, 2015) runs at rate 0.03 for 1500 iterations.

AuctionNet. Prices are discretized to the shared price grid $[K]$ with $K = 300$. Training uses 411,305 rows sampled from the benchmark's advertising periods 7–8. Evaluation uses 400,000 held-out rows from periods 9–10, drawn balanced as 200,000 wins and 200,000 losses. Calibration metrics down-weight the win rows so the effective composition matches the natural win rate of 0.0735, measured on all positive-bid rows of period 9. The backbone is an embedding MLP over seven categorical features. The features are advertiser, category, time step, and the binned budget, CPA constraint, and the benchmark's value prediction fields pValue and pValueSigma.

iPinYou. Prices come from 417,196 won impressions of iPinYou campaign 2259 in the processed DLF data (Ren et al., 2019). Every retained row is a win at a nearly constant logged bid, so the logged bids cannot censor. We keep the real prices and reassign the simulated bid stream shipped with the DLF data across rows by a rank mixture. With $r(z_i)$ the normalized rank of the price z_i and u_i i.i.d. uniform noise, row i receives the bid whose rank matches the rank of $m_i = \lambda\, r(z_i) + (1-\lambda)\, u_i$ for each $\lambda \in \{0, 0.25, 0.5, 0.75\}$. Since the price rows are untouched and the bids are permuted, both margins are fixed. At $\lambda = 0$ the pairing is independent. At $\lambda = 1$ the k-th highest bid would meet the k-th highest price, a comonotone coupling outside the positive-density hypothesis of Theorem 1, so the grid stops at 0.75. Aligning the ranks drives the censoring rate from 0.76 at $\lambda = 0$ to 0.92 at $\lambda = 0.75$. The estimator without features is fit on all coupled rows, since the true margin is computed from those rows. The proposed model uses the sixteen bag-of-ids features of the DLF release in one shared embedding. A seeded 20% row split provides the held-out CRPS, while the marginal mixture $n^{-1} \sum_i \hat{F}(k \mid x_i)$ is computed on every row against the true margin.

AuctionGym. We reimplement the AuctionGym simulator (Jeunen et al., 2023) with 200,000 auction rounds, four competitors, twelve items, and five-dimensional context embeddings. The coupling grid and the fit on all rows match the iPinYou benchmark. Each round draws one context vector shared by every agent, who bids the maximum of predicted click-through rate times item value over its own catalog. The market price is the highest competing bid. Here the rank mixture ties the bid of the tracked bidder, the agent whose log we analyze, to that bidder's own value rather than to the price, reassigning the simulator's truthful bids from independence at $\lambda = 0$ to the value's exact ranking at $\lambda = 1$. The price is linked to that value only through the shared context, so the censoring rate stays nearly flat, from 0.62 to 0.66 across the grid.

Metrics. The CDF errors in the tables average the absolute margin deviation over price buckets up to the 90th percentile of the bids, the bid range marked by the dotted boundary in Figure 1 (right). Above that line almost every auction is censored, so no method has data there. The CRPS of a fitted CDF at a realized price is the sum over the price grid of the squared gap between that CDF and the unit step at the price (Gneiting & Raftery, 2007). The win-probability and surplus errors of Table 3 on iPinYou are computed at the bids in that range. Win-probability error is the mean absolute error of $\hat{F}(b-1)$ over those bids. The surplus at bid b is $S(b) = \sum_{k<b} F(k) = \mathbb{E}[(b-Z)^+]$, the expected surplus of a bidder with value b who pays the market price on a win. This is the second-price form. A first-price bidder with value v pays its own bid, so its expected surplus is $v - b$ times its win probability, a different functional of the same F that the table does not report. Relative surplus error is the mean absolute error of $\hat{S}(b) = \sum_{k<b} \hat{F}(k)$ over the same bids divided by their mean true surplus.

B.3 Comparison estimators

The covariate column. The covariate network takes the bid as one more embedded id on AuctionNet. On iPinYou and AuctionGym, we also try a continuous input and report the better form, the id on iPinYou and the continuous input on AuctionGym. On AuctionGym the continuous form still misses the margin by 0.303 ± 0.042 at independence and 0.291 ± 0.046 at $\lambda = 0.75$, an error that does not move with the coupling. Because every tail is unconstrained beyond its bid, the main table leaves those cells empty. On AuctionNet its CRPS matches the censored network. With the time step withheld it trails even the product-limit (Figure 2, right).

The bid-landscape baseline. The DLF model of Section 4 is a GRU over the price grid with hidden width 64. Its step-k output is the conditional hazard $P(z = k \mid z \geq k, x)$ and its initial state is the id embedding shared with the other iPinYou networks. Features, split, censored likelihood,

Table 4: Dependence ablation. Each row fixes how the dependence enters the fit. AuctionNet reports the CRPS of the proposed model. iPinYou at $\lambda = 0.5$ reports the CDF error of the copula-graphic recursion with the dependence supplied, or of the fits with the dependence estimated.

Dependence in the fit	AuctionNet CRPS	iPinYou CDF error
None, censored likelihood	0.052±0.004	0.058±0.006
None, θ frozen at 0	0.049±0.003	0.059±0.006
None, DLF	0.048±0.001	0.060±0.006
None, copula-graphic recursion	–	0.074
Half the true correlation, copula-graphic recursion	–	0.052
True Kendall dependence, Gaussian recursion	–	0.007
Estimated, DCSurvival	0.045±0.003	0.060±0.000
Estimated (proposed)	0.042±0.002	0.009±0.003
Oracle network	0.039±0.001	0.004±0.003

and optimizer budget match the other iPinYou networks. Beyond the $\lambda = 0.5$ column of Table 4, its CDF error is 0.010 at $\lambda = 0$ and 0.114 at $\lambda = 0.75$, against 0.117 for the censored network and 0.075 for the proposed model there.

The DCSurvival baseline. The DCSurvival fit of Section 4 keeps only $(\min(Z, B), W)$ per auction, the standard right-censoring format. Both margins are then learned jointly with the copula. The generator is a completely monotone exponential mixture with ten components, the class of Zhang et al. (2024) written on the discrete grid. The family includes the independence copula $C(u, v) = uv$ (Zhang et al., 2024) and so contains the truth at $\lambda = 0$. Wins at price k contribute the exact cell $P(Z = k,\, B > k)$ and losses at bid k the cell $P(Z \geq k,\, B = k)$. On the iPinYou datasets ($z < b$ wins) the margins are hazard vectors, fit full batch for 3000 Adam iterations at rate 0.03. On AuctionNet they are two backbones under the shared protocol, the generator again at rate 0.03. The fitted Kendall dependence is 0.017 ± 0.001 on AuctionNet and 0.61 on the $\lambda = 0.5$ dataset against its planted 0.51. At $\lambda = 0$ it returns dependence 0.50 with a margin off by 0.122, where the product-limit misses by 0.004. This is not an optimization failure. Its average negative log-likelihood matches the truth's to within 1.1×10^{-4}. DCSurvival uses records in which the censoring value is kept only when it censors, leaving identification to restrictions on the margins. The auction record keeps the bid on wins as well, which lets Theorem 1 identify θ from the win-region cells with the price margin left nonparametric. On AuctionNet, where the fitted dependence is small, its CRPS of 0.045 does improve on the censored network (Table 4).

Dependence as a function of features. Replacing the shared scalar with $\theta(x) = 0.99 \tanh h(x)$, where h is a bag-of-ids network, does not improve the fit on iPinYou at $\lambda = 0.5$ or 0.75. The CRPS of $F(k \mid x)$ stays within 0.005 of the shared scalar, the bid-conditional CRPS worsens by 0.008 to 0.023, and the margin error moves both ways, up at $\lambda = 0.5$ and down at 0.75. The fitted $\theta(x)$ spans -0.6 to the 0.99 cap between its tenth and ninetieth percentiles, with its mean at the shared value, so the likelihood trades it against $F(k \mid x)$ where the win cells constrain it weakly. The shared scalar, a restriction like the Gaussian family, lets a feature group whose smallest bid never wins take its dependence from groups where it does.

B.4 Experiment constructions

Withheld-signal construction. The right panel of Figure 2 isolates endogeneity from a bidder signal the estimator never sees. Two predictors of log price are fitted out of fold on the AuctionNet training rows, one on the six visible features and one on all seven. The withheld feature is the time step. The other six features explain almost none of the price variance. Write m_{obs} and m_{full} for the two predictors and x_{obs} for the visible features. The bidding rule scores each auction by

$$s = m_{\text{obs}}(x_{\text{obs}}) + \gamma \left(m_{\text{full}}(x) - m_{\text{obs}}(x_{\text{obs}}) \right) + \sigma\, \varepsilon, \qquad \gamma \in \{0, 0.5, 1\},$$

with ε standard normal and σ half the standard deviation of m_{full}. Each auction receives a logged bid by the rank of s, so the bid margin is the logged one at every γ. Every estimator receives x_{obs} only, the covariate network taking the logged bid as one more embedded id. Training follows the

shared protocol at whole epochs up to 1500 updates, since the construction is a coupling grid rather than the fixed benchmark of Table 1. Censoring rises from 0.90 at $\gamma = 0$ to 0.96 at $\gamma = 1$ as the bids align with the prices. The fitted dependence $\hat{\theta}$, averaged over the ten seeds, is -0.06 at $\gamma = 0$ and 0.47 at both $\gamma = 0.5$ and $\gamma = 1$, where the rank correlation of bid and price is 0.48 and 0.61. The construction plants no Gaussian copula, so no single true θ exists; at $\gamma = 1$ the fit stops short of the rank correlation and still recovers 93% of the censored-to-oracle gap.

Planted couplings. The planted rows of Table 2 draw pairs from the named copula at Kendall dependence 0.51, the value of the $\lambda = 0.5$ dataset. The price side of each pair is rank-matched to the observed prices while the bid side reassigns the DLF bid stream by rank, leaving the bid margin unchanged. Censoring is 0.87 on the Clayton dataset and 0.85 on the Gumbel one. On these rows likelihood selection chooses among the Gaussian, Clayton, Gumbel, and Frank families, and the copula-graphic recursion of Table 2 receives the planted family at the same dependence.

Signal-driven couplings. The last two rows of Table 2 run at a signal share $w \in \{0, 0.25, 0.5, 0.75, 1\}$, with the $w = 1$ rows reported. On AuctionGym only the tracked bidder changes, computing its value from $\sqrt{w}\, c + \sqrt{1-w}\, \xi$ in place of the shared context c, with ξ a private standard normal vector. Since c and ξ are independent standard normal, this input is standard normal at every w. The bid margin therefore stays fixed across the grid and only its link to the price changes. On iPinYou a predictor of log price is fitted out of fold on the sixteen bag-of-ids features by two-fold target encoding and a ridge combination, reaching an out-of-fold R^2 of 0.45. The DLF bids are reassigned by the rank of $\sqrt{w}\, \tilde{m}(x) + \sqrt{1-w}\, \varepsilon$, with $\tilde{m}$ the standardized prediction, keeping the bid margin. Given all sixteen features, this coupling is exogenous and the proposed model would face no endogeneity, so the table reports the estimator without features. Across the shares, censoring rises from 0.63 to 0.67 on AuctionGym and from 0.76 to 0.85 on iPinYou, while the Kendall dependence rises from zero to the values in Table 2. At every positive share, the Gaussian fit removes 73 to 88% of the product-limit error. The copula-graphic recursion receives the Gaussian family at the measured τ through $\rho = \sin(\pi\tau/2)$.

Verifying the family. Per bid bucket, the fitted model implies a multinomial over the win prices and the loss mass. We group adjacent cells to an expected count of five and compare Pearson's X^2 with its degrees of freedom, counting only the hazards the likelihood constrains. On the planted Clayton dataset, the statistic of the Gaussian fit is 42 times its degrees of freedom while the Clayton fit sits at 1.0. On a Gaussian-coupled control dataset at the same strength, the Gaussian fit sits at 1.0 and the Clayton fit at 6.6. Across 100 replications at $n = 50{,}000$ drawn from either family with the iPinYou margins, the correct-family statistic stays below 1.3 while the wrong-family statistic stays above 2.8. On the coupling grid of Table 1, the Gaussian fit sits at 1.9 on AuctionGym at $\lambda = 0.75$. On iPinYou at $\lambda = 0.5$, every family lies between 14 and 24.

Fixing the dependence. Setting the Gaussian dependence to $\rho_\star = \sin(\pi\hat{\tau}/2)$, with $\hat{\tau}$ Kendall's τ of the uncensored pairs, gives the estimator without features a common reference across families. At $\rho_\star$ the Pearson ratio X^2/df is 27.2 on the planted Clayton dataset, 4.9 on the planted Gumbel dataset, and 19.3 on iPinYou at $\lambda = 0.5$, with CDF errors of 0.021, 0.007, and 0.007 against naive errors of 0.096, 0.075, and 0.081. The free fits imply τ values differing from $\hat{\tau}$ by $+0.07$, -0.08, and -0.04. Their negative log likelihoods are lower than those of the fixed fits on all three datasets, while their CDF errors are higher by 0.002, 0.009, and 0.009. The ratio falls on Gumbel and iPinYou and rises on Clayton, so it does not track CDF error monotonically. On the Gaussian control, fixing ρ at $\rho_\star - 0.1$ or $\rho_\star + 0.1$ raises the ratio from 1.0 to 2.8 and 11.3, respectively, with CDF error increasing from 0.001 to 0.023 and 0.037.

Bid replay. The replay holds out a seeded 20% of the AuctionNet training rows and trains the estimators on the rest. An estimator bids the smallest b with $\hat{F}(b-1 \mid x) \geq 0.1$, the win probability at bid b under the log's strict win rule. The replay then records the realized win rate over the held-out rows. At 5% and 20% targets the ordering is the same, the corrected distribution landing 5 to 11 pp above target, the censored network 12 to 24 pp, and the oracle within 3 pp. Training and replay rows come from the same periods, since the price law shifts into periods 9 and 10. An out-of-period replay would measure that drift rather than the correction.

B.5 Rate and coverage under correct specification

Two synthetic designs are used. In the first there are no features. In the second the feature takes six values with unequal probabilities. Price and bid margins are Beta-shaped on a grid of 40 prices, one pair per feature value, and one latent Gaussian coupling with $\theta_0 = 0.6$ is shared. The first design is fitted by the estimator without features and the second by the saturated table of Proposition A.5. Over five sample sizes from 12,500 to 200,000 the sup error falls with log-log slope -0.61 in the first design and -0.52 in the second. The error of $\hat{\theta}$ falls with slope -0.49 and -0.52. Over 100 replications at $n = 50{,}000$ the 95% two-stage sandwich intervals cover θ_0 at 0.98 in both cases and a mid-support CDF value at 0.95 and 0.98, slightly conservative since the standard deviation of the estimates is 0.87 to 1.04 times the mean standard error. Standard errors from the stage-two likelihood alone, with the bid margin held at its empirical value, are 4% smaller for θ and 6 to 8% smaller for the CDF value, so the first-stage term of Theorem 2 is needed.

B.6 Fixed-bid records in finite samples

We check Proposition A.3 on semi-synthetic streams with the true price known. A 5% share of the auctions is rebid at a fixed bid t, the 0.8 quantile of the bid margin, with the win and the recorded price recomputed from the true price. The fixed-bid estimate follows the proof. The marked auctions give F up to t as a win frequency. Above t, the recursion of Theorem 1 runs on the unmarked auctions, with θ taken from the equation at t.

Table 5 uses six streams. The Gaussian stream couples the iPinYou margins at Kendall $\tau = 0.51$. Removing its anchor moves every price at or below $b_0 + 8$ upward, so that $F(b_0) = 0$, as AuctionGym at $\lambda = 0.5$ already satisfies. The planted streams are those of Table 2 and are fitted in their true family. The other streams are fitted in the Gaussian.

Table 5: Fixed-bid records. Mean absolute CDF error above the fixed bid t, averaged over twenty draws of the marked share. The product-limit assumes independence. The Gaussian fit is the estimator without features, anchored at the smallest bid and fitted without fixed-bid records. The last column is the construction of Proposition A.3 at a marked share of 5%.

Stream	Product-limit	Gaussian fit	Fixed bid, 5%
Gaussian	0.188	0.001	0.002
Gaussian, anchor removed	0.217	0.004	0.002
AuctionGym $\lambda = 0.5$, $F(b_0) = 0$	0.052	0.035	0.007
Clayton planted	0.202	0.092	0.002
Gumbel planted	0.186	0.053	0.002
iPinYou $\lambda = 0.5$	0.197	0.061	0.020

With the family right, the fixed-bid records cost little. On the Gaussian stream, the fixed-bid estimate lands within 0.002 of the Gaussian fit, and removing the anchor leaves it unchanged, since it never uses b_0. Where the Gaussian family is wrong, it closes most of the remaining error, reaching 0.007 against 0.035 on AuctionGym and 0.002 against 0.05 to 0.09 on the planted streams.

The roots at $s \leq t$ give the family check of Proposition A.3. We take roots at the bid quantiles below t with at least twenty marked wins. Two statistics, the spread of the roots and a Wald statistic on the overidentifying equations, are calibrated by a stratified bootstrap of the log with 200 draws. On 200 fresh draws of each coupling on the iPinYou margins, the Gaussian fit is rejected at the 5% level on every Clayton and Gumbel draw, while the true family is rejected in 3.5 to 8.5% of draws. On the iPinYou rank mixture, the Gaussian equation at t has no root in any draw, which counts as a rejection, and its row in Table 5 takes θ at the end of its bracket.

B.7 The band under affiliation

We compute the bounds of Theorem 1(i) in two settings, averaging over prices up to the 0.9 quantile of the bid. On the population grid, Gaussian couplings of the iPinYou margins at Kendall τ of 0.16, 0.33, and 0.54 give an affiliation band of average width 0.016, 0.039, and 0.070, against 0.33 for the band of Proposition 1. The product-limit lies below the band, by 0.012, 0.024, and 0.037 on average.

On the AuctionNet training log, we round each bid down to one of twenty quantile points. This is a valid log of lower bids and stays affiliated, by the argument of Remark A.1. From the log alone, the product-limit falls below the lower edge by up to 0.043, with a one-sided 95% bootstrap bound of 0.040 at that price (0.005 and 0.054 with ten and forty points). The true price confirms $F \geq F_L$ up to 10^{-4}. We omit the upper edge, whose plug-in rests on minima of cell ratios and is unstable at this size.

C Additional related work

An estimator still needs a guarantee on how fast it converges and how its error is distributed. Guo et al. (2023) establish root-n normality with a sandwich covariance for semiparametric estimators under the conditional independence restrictions of a criss-cross missingness model. In survival analysis, Chen (2019) derive finite-sample bounds under conditionally independent censoring. With every bid observed and the threshold rule known, Theorem 2 gives joint root-n normality for the price margin on the bid support and the copula dependence parameter. Its sandwich covariance includes the first-stage error from estimating the bid margin. El Balghiti et al. (2019) connect prediction error to downstream optimization with decision-focused generalization bounds, complementing our analysis at the decision layer. We identify and estimate the price margin supplied to the bidding rule, then measure its effect on win rates, surplus, and win-probability calibration.