Analysis Report

Simulation Results for Informal Bid Admission Cutoff Estimation

Austin Li

Contents

1 Executive Summary

This report presents the results from the simulation exercise designed to validate the threshold/interval framework for estimating admission cutoffs in M&A-style auctions. The exercise was conducted in response to the December 18, 2025 meeting notes, which outlined two simulation tasks:

- **Task A**: Estimate a single constant cutoff b^{I*} when only admission/rejection decisions are observed
- **Task B**: Estimate type-specific cutoffs b_S^{I*} and b_F^{I*} for two bidder types (S and F)

1.1 Key Findings

Observed-sample definition (formal-stage conditioning). In this simulation exercise, N denotes the number of auctions that *reach the formal stage* (i.e., at least one bidder is admitted). Auctions with zero admitted bidders ("all-reject") are generated but treated as unobserved and excluded from the returned sample. Auctions where everyone is admitted ("all-admit") are retained via a *one-sided upper bound* on b^{I*}.

Main findings (based on 10 replications per N).

1. **Task A (Single Cutoff): baseline recovery but undercoverage at large N.** In the baseline run with $N = 100$ observed auctions, the posterior mean is $\hat{\mu} = 1.387$ (true $b^* = 1.400$) with a 95% credible interval [1.367, 1.406] (coverage achieved in this run). In sample-size sensitivity, the bias remains small but negative and *coverage deteriorates sharply as N grows* (e.g., 80% at $N = 100$, 60% at $N = 200$, and 0% at $N = 500$ in this run), consistent with the estimator becoming precise around the wrong value under formal-stage conditioning.

2. **Task B (Type-Specific Cutoffs): ordering is learnable, but coverage for μ_F weakens with N.** The baseline run recovers the ordering $b_S^* > b_F^*$ with $P(\mu_S > \mu_F \mid \text{data}) = 1.00$ and tight posteriors. In sensitivity analysis, $P(\mu_S > \mu_F)$ is already high at $N = 20$ (about 92%), but the 95% coverage rate for type F falls below nominal as N increases (e.g., 50% at $N = 200$ in this run).

3. **Incomplete auctions among observed deals are rare under this DGP.** Conditional on reaching the formal stage, the main incomplete case is "all-admit" (one-sided upper bounds). In the baseline parameters used here, this occurs in roughly 3–5% of *observed* auctions.

2 Methodology Recap

2.1 The Threshold/Interval Framework

The estimation approach treats admission as governed by a deterministic screening rule:

$$j \in \mathcal{A}_i \iff b^I_{ij} \geq b^{I*}_i$$

where b^I_{ij} is bidder j's informal bid and b^{I*}_i is the latent admission threshold.

Observed admit/reject decisions imply bounds on the threshold. Define

$$L_i := \begin{cases} \max_{j \in \mathcal{R}_i} b^I_{ij}, & \mathcal{R}_i \neq \emptyset \\ -\infty, & \mathcal{R}_i = \emptyset \end{cases} \qquad U_i := \begin{cases} \min_{j \in \mathcal{A}_i} b^I_{ij}, & \mathcal{A}_i \neq \emptyset \\ \infty, & \mathcal{A}_i = \emptyset \end{cases}$$

so that $b^{I*}_i \in [L_i, U_i]$.

Formal-stage conditioning. In the simulations (and in the intended empirical setting), the observed dataset is conditional on reaching the formal stage. For Task A, this means we observe only auctions with at least one admitted bidder, i.e. $U_i < \infty$; auctions with $\mathcal{A}_i = \emptyset$ ("all-reject", $U_i = \infty$) are treated as unobserved and excluded. Auctions with $\mathcal{R}_i = \emptyset$ ("all-admit", $L_i = -\infty$) are retained and contribute a one-sided upper bound.

2.2 MCMC Data Augmentation

The estimation uses Gibbs sampling with data augmentation:

$$b^{I*}_i = \mu + \nu_i \tag{1}$$

$$\nu_i \sim \mathcal{N}(0, \sigma^2), \quad \nu_i \in [L_i - \mu, U_i - \mu] \tag{2}$$

Each MCMC iteration:

1. Samples latent ν_i from truncated normals (data augmentation step)
2. Updates μ via conjugate normal posterior
3. Updates σ^2 via conjugate inverse-gamma posterior

2.3 Simulation Parameters

Baseline DGP (Task A):

- $N = 100$ *observed (formal-stage)* auctions, $J = 3$ bidders per auction
- Valuations: $v_{ij} = 1.3 + \epsilon_{ij}$, $\epsilon_{ij} \sim \mathcal{N}(0, 0.2^2)$
- True cutoff: $b^{I*} = 1.4$

Baseline DGP (Task B):

- Same valuation structure
- $N = 100$ *observed (formal-stage)* auctions, $J = 3$ bidders per auction
- Type-specific cutoffs: $b_S^{I*} = 1.45$, $b_F^{I*} = 1.35$
- Type probability: $P(\text{Type S}) = 0.5$

3 Task A Results: Single Constant Cutoff

3.1 Baseline Simulation

Figure 1 presents the MCMC diagnostics for the Task A baseline simulation with $N = 100$ auctions.

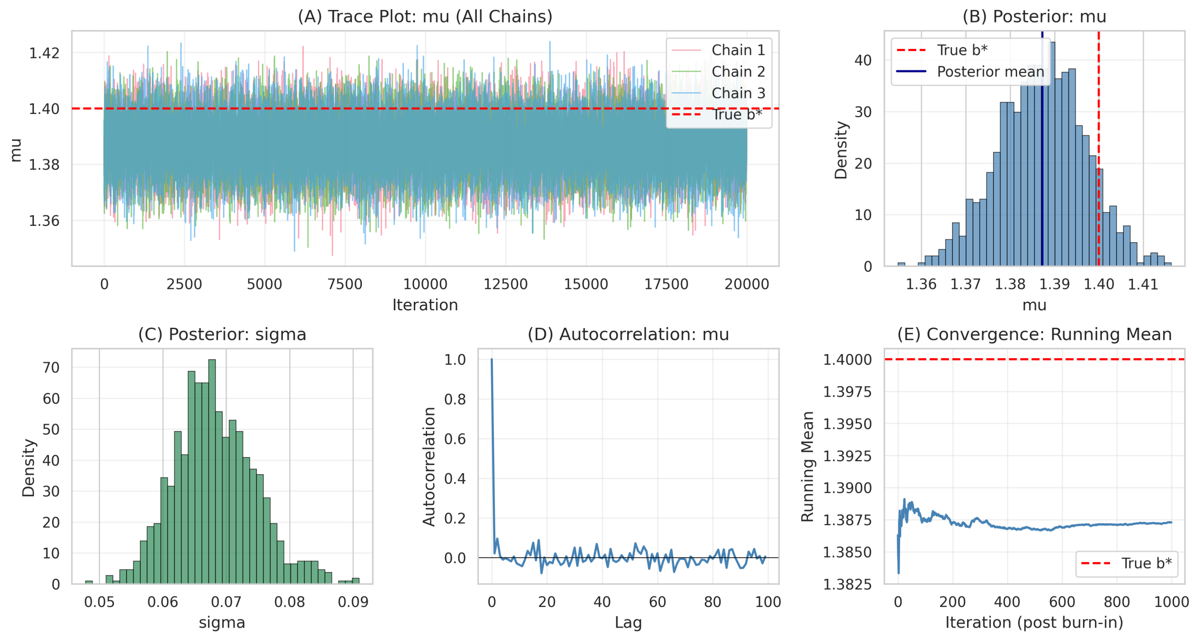

Figure 1: Task A MCMC Diagnostics. Left panels show trace plots for μ and σ across three chains (good mixing indicated by overlapping chains). Right panels show posterior histograms with the true value (dashed red line) and posterior mean (solid blue line).

Observations:

- Chains mix well and converge (Gelman–Rubin $\hat{R} \approx 1.00$).
- Baseline estimate: posterior mean $\hat{\mu} = 1.387$ with 95% credible interval [1.367, 1.406] (true $b^* = 1.400$).
- The estimate is slightly below the true cutoff in this run, consistent with the small negative bias seen in the sensitivity analysis under formal-stage conditioning.

Figure 2 visualizes the *two-sided* interval constraints $[L_i, U_i]$ for auctions with both admitted and rejected bidders ("complete" auctions). One-sided all-admit auctions are retained in estimation but omitted from this figure for readability.

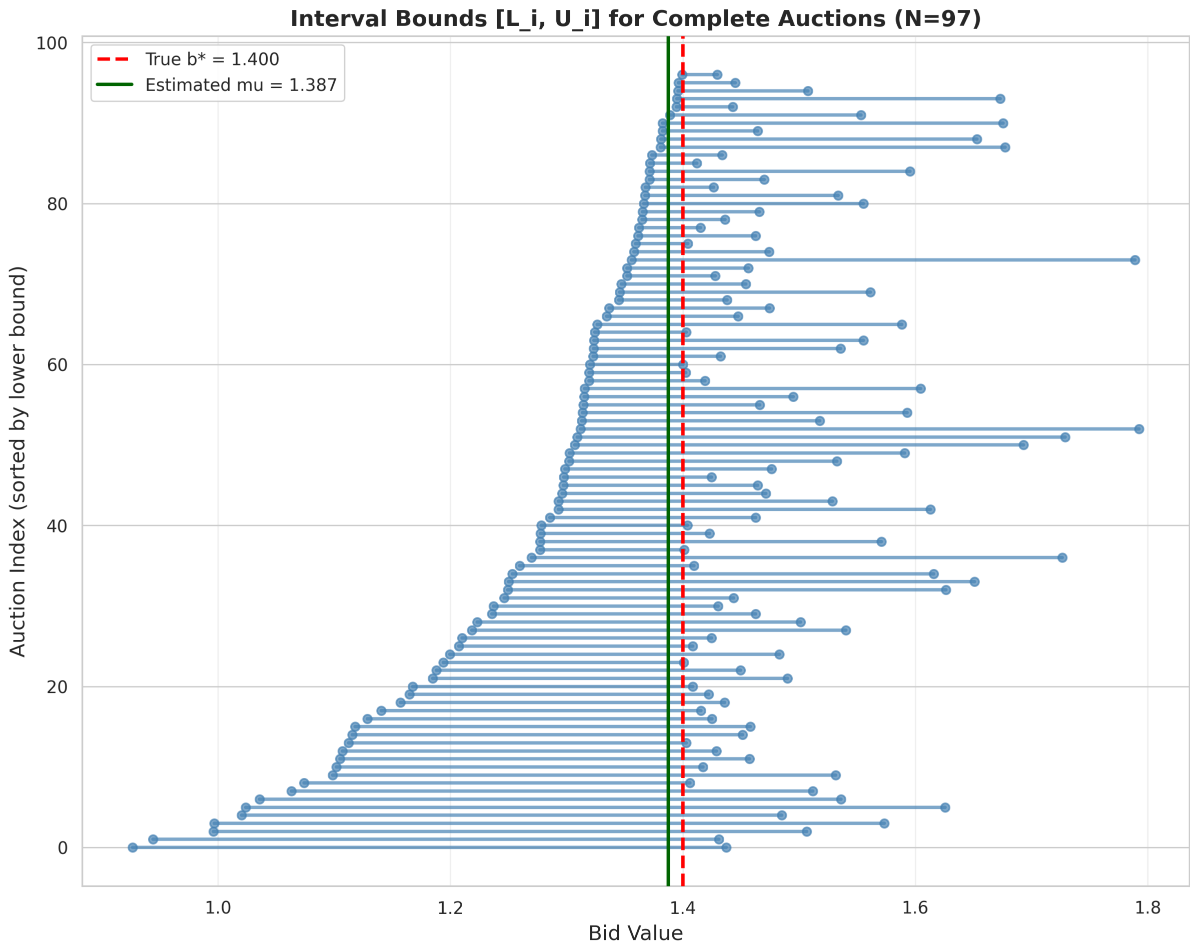

Figure 2: Task A Interval Bounds. Each horizontal line represents one auction's interval $[L_i, U_i]$. The true cutoff (red dashed) and estimated cutoff (blue dashed) are shown. Auctions providing tighter intervals contribute more information.

Observations:

- Two-sided intervals cluster around the true cutoff, with meaningful variation in width across auctions.
- The posterior mean lies inside most two-sided intervals, reflecting how the likelihood aggregates the auction-level bounds.

3.2 Sensitivity Analysis: Sample Size

A key concern is performance with the small sample sizes typical of hand-collected M&A data. Figure 3 shows how estimation quality varies with N.

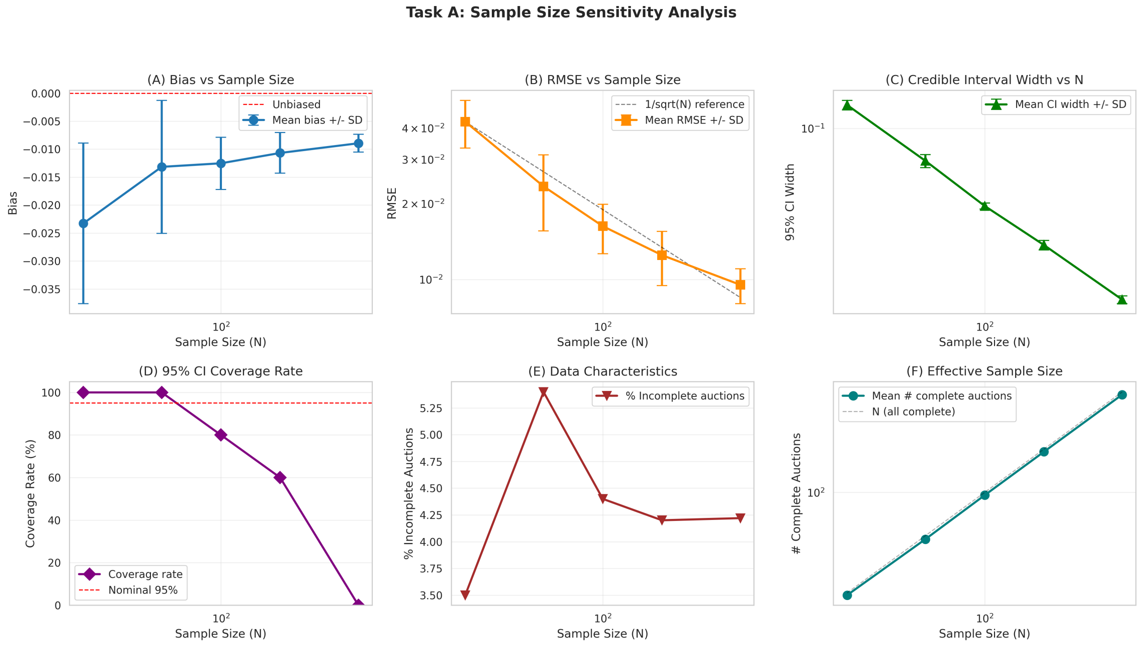

Figure 3: Task A sensitivity to sample size under formal-stage conditioning. Results across $N \in \{20, 50, 100, 200, 500\}$ with 10 replications each. Bias is small but negative, credible intervals shrink with N, and empirical coverage falls substantially below the nominal 95% level for large N in this run.

Key findings:

Table 1: Task A Performance Metrics by Sample Size

N	Mean Bias	Mean RMSE	Mean CI Width	Coverage	% Incomplete
20	-0.0233	0.0424	0.1328	100%	3.5%
50	-0.0132	0.0235	0.0677	100%	5.4%
100	-0.0125	0.0163	0.0393	80%	4.4%
200	-0.0107	0.0125	0.0245	60%	4.2%
500	-0.0089	0.0095	0.0127	0%	4.2%

Interpretation:

- **Bias is small but persistently negative** under formal-stage conditioning (about -0.009 to -0.023 across N in this run).
- **Credible intervals shrink quickly with** N, but **empirical coverage falls** as N increases (down to 0% at $N = 500$ in this run), indicating undercoverage relative to the nominal 95% target.
- At $N \approx 20$ (the expected hand-collected sample size), intervals are wide enough that coverage can look good, but this should not be interpreted as evidence of correct calibration once the sample-selection mechanism is accounted for.

3.3 Incomplete Auctions

With formal-stage conditioning, "all-reject" auctions ($\mathcal{A}_i = \emptyset$) are treated as unobserved and excluded from the observed sample. Among observed auctions, the relevant incomplete case is "all-admit" ($\mathcal{R}_i = \emptyset$), which provides only a one-sided upper bound ($L_i = -\infty$).

In the baseline run with $N = 100$ observed auctions, 3% of observed auctions are all-admit. To obtain $N = 100$ observed auctions, the simulator initiates more auctions and drops unobserved all-reject cases (e.g., 155 initiated vs. 100 observed in the baseline run).

4 Task B Results: Type-Specific Cutoffs

4.1 Baseline Simulation

Task B estimates separate cutoffs for Type S ($b_S^{I*} = 1.45$) and Type F ($b_F^{I*} = 1.35$) bidders. Figure 4 shows the MCMC diagnostics.

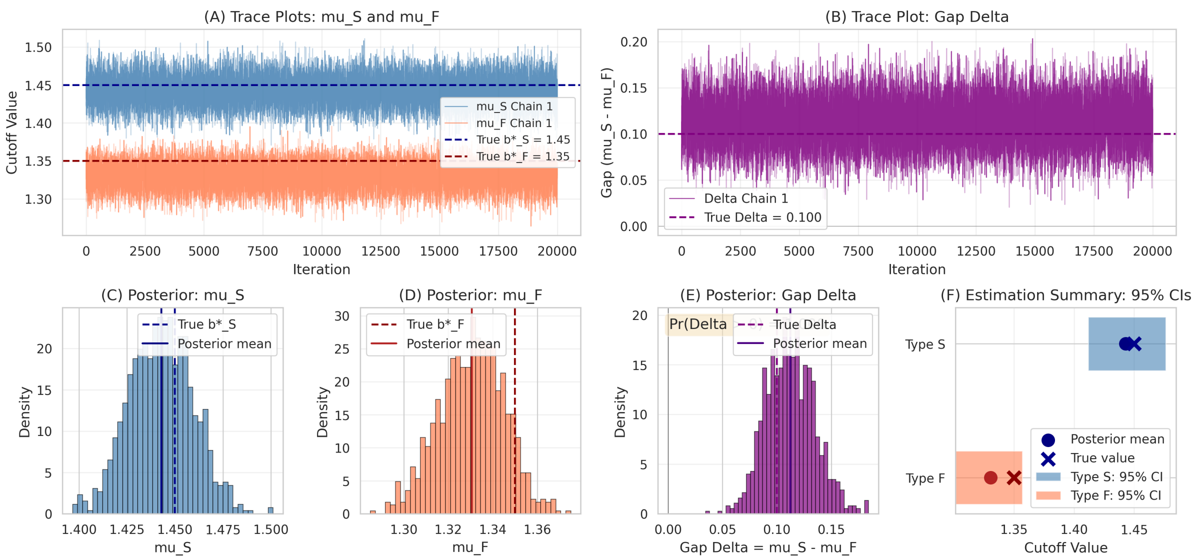

Figure 4: Task B MCMC diagnostics for μ_S, μ_F, and the gap $\Delta = \mu_S - \mu_F$ (three chains). Dashed reference lines indicate true values; solid reference lines indicate posterior means.

Observations:

- Chains mix well and converge ($\hat{R} \approx 1.00$ for both types).
- Baseline estimates: $\hat{\mu}_S = 1.443$ with 95% CI [1.412, 1.476] (true 1.450) and $\hat{\mu}_F = 1.331$ with 95% CI [1.301, 1.357] (true 1.350).
- Effective sample sizes differ by type because some auctions provide no information for a given type (e.g., in the baseline run the sampler uses 86 auctions with S information and 91 with F information out of $N = 100$ observed auctions).

Figure 5 shows the type-specific *two-sided* interval constraints (auctions with both admitted and rejected bidders of that type). One-sided type-specific bounds are used in estimation but omitted from this figure.

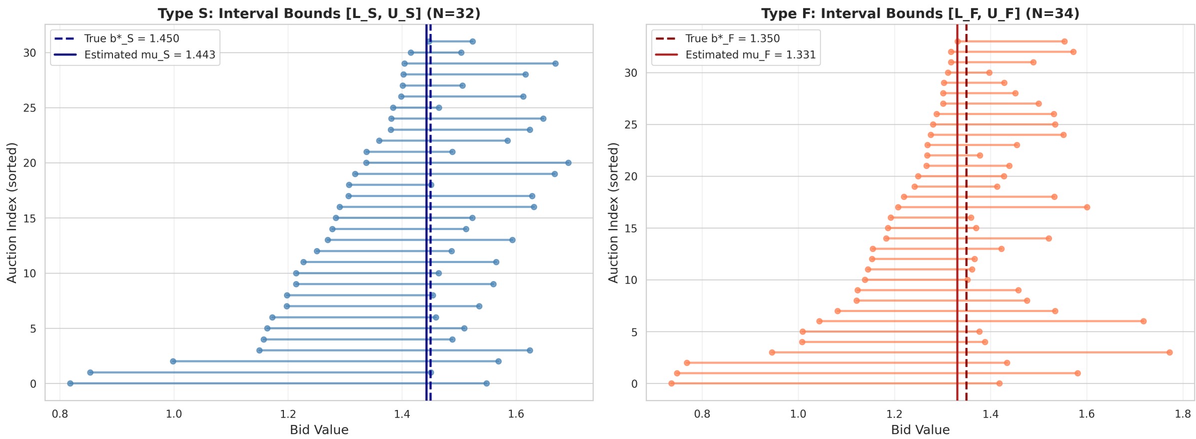

Figure 5: Task B type-specific two-sided intervals. Left: Type S intervals. Right: Type F intervals. True cutoffs (dashed) and posterior means (solid) are shown for each type.

4.2 Cutoff Gap Estimation

A central question is whether the data can distinguish between type-specific cutoffs. The posterior for the gap $\Delta = \mu_S - \mu_F$ provides this inference.

Key results:

- True gap: $\Delta^* = 1.45 - 1.35 = 0.10$
- Estimated gap (baseline run): posterior mean $\hat{\Delta} = 0.112$
- $P(\mu_S > \mu_F \mid \text{data}) = 1.00$ in the baseline run

This suggests that, under this DGP, **moderate sample sizes can identify economically meaningful differences** in how the target screens different bidder types.

4.3 Sensitivity Analysis: Sample Size

Figure 6 examines Task B performance across sample sizes.

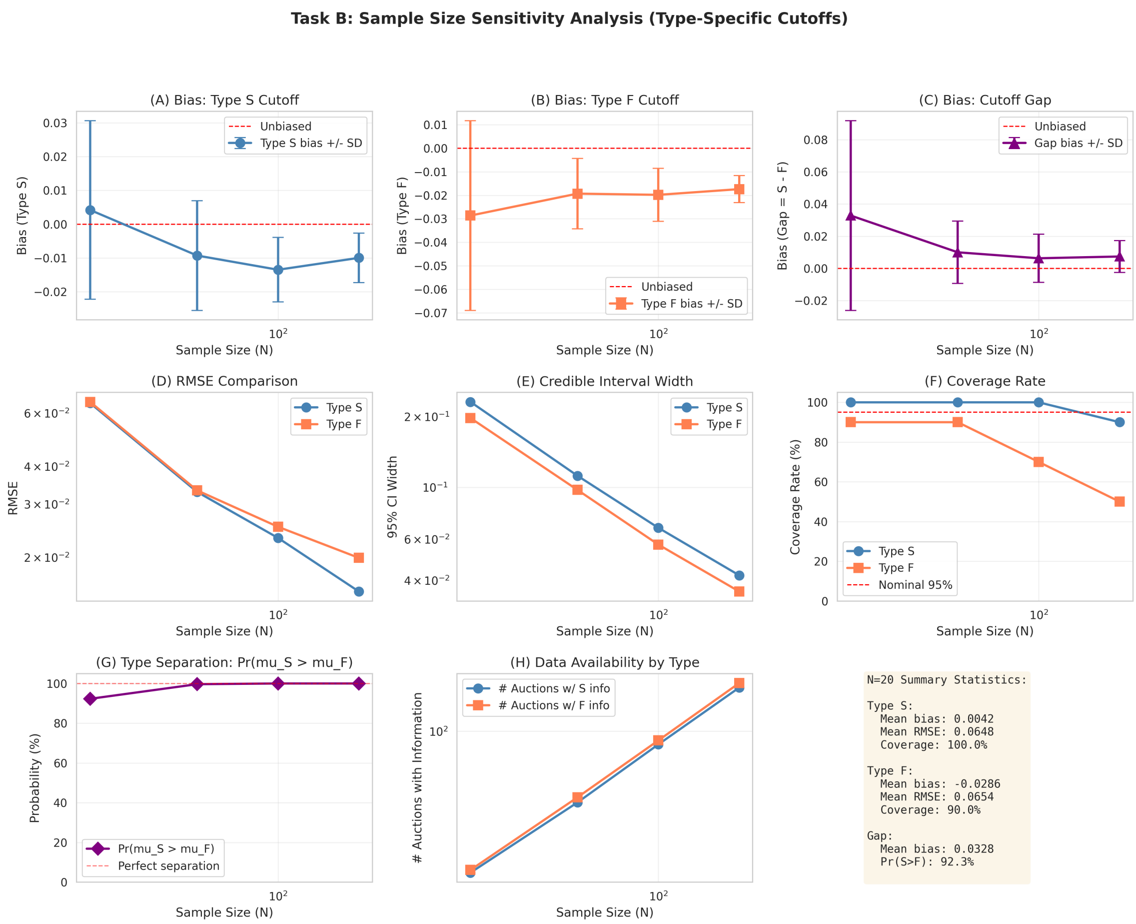

Figure 6: Task B sensitivity to sample size under formal-stage conditioning. RMSE and credible interval widths fall with N, $P(\mu_S > \mu_F)$ increases toward 1, and empirical coverage for type F falls well below 95% at larger N in this run.

Key findings for $N = 20$:

Table 2: Task B Performance at $N = 20$ (Critical for Real Data)

Metric	Type S	Type F
Mean Bias	0.0042	-0.0286
Mean RMSE	0.0648	0.0654
Coverage	100%	90%
$P(\mu_S > \mu_F)$	92.3%	

Interpretation:

- At $N \approx 20$, the estimator already assigns high probability to the correct ordering ($P(\mu_S > \mu_F) \approx 92\%$).
- Uncertainty is substantial at small N (wide credible intervals), which can yield high coverage in small samples.

- Coverage for type F deteriorates as N grows in this run, suggesting the same selection/truncation issue as in Task A when the likelihood does not explicitly condition on reaching the formal stage.

5 Key Findings and Takeaways

This section directly addresses the deliverables requested in the December 18, 2025 meeting notes.

5.1 Recovery Accuracy for Cutoffs

Finding 1: Baseline recovery is reasonable, but calibration depends on modeling formal-stage selection.

Both tasks produce well-behaved MCMC output and plausible posterior summaries in baseline runs. However, under formal-stage conditioning (dropping all-reject auctions as unobserved), the current likelihood does not explicitly condition on being observed. In the sample-size sensitivity runs, this manifests as **undercoverage at larger** N (Task A: 80% at $N = 100$ and 0% at $N = 500$; Task B: type-F coverage 50% at $N = 200$ in this run), despite small average bias.

5.2 Frequency of Informative Intervals

Finding 2: Among observed auctions, two-sided bounds are common, but selection changes what is observed.

For Task A, conditional on reaching the formal stage, most observed auctions provide two-sided bounds (about 95% complete in the sensitivity runs at $N = 100$). The remaining observed auctions are typically all-admit, contributing one-sided upper bounds. All-reject auctions are not part of the observed sample by design, but their absence is informative about the selection mechanism.

For Task B, most observed auctions provide information for each type (e.g., at $N = 20$ the sensitivity runs average 17.5 auctions with S information and 18.1 with F information), while two-sided bounds *by type* are less frequent (and the interval figure displays only those two-sided-by-type auctions).

5.3 Sensitivity to Cutoff Magnitude and Type Separation

Finding 3: Type separation is learnable in this DGP; selection remains the main concern.

In the current parameterization ($b^*_S - b^*_F = 0.10$), Task B strongly distinguishes types as N increases and already delivers a strong ordering signal at $N \approx 20$ (about 92% in this run). The main limitation highlighted by the sensitivity exercises is not identification of the ordering, but **credible-interval calibration under formal-stage conditioning**.

5.4 Performance at $N \approx 20$

Finding 4: At $N \approx 20$, the method is feasible and informative, but uncertainty is substantial.

At small N, credible intervals are wide and can exhibit high empirical coverage in these runs. For Task B, even at $N \approx 20$, the posterior often favors $\mu_S > \mu_F$. These small-sample results are consistent with feasibility for a hand-collected sample, while reinforcing that formal-stage selection needs to be modeled before interpreting credible intervals as calibrated uncertainty.

6 Conclusion

The simulation exercise validates the *mechanics* of the threshold/interval approach: interval construction is straightforward, the Gibbs samplers converge reliably, and the framework can recover type ordering in Task B at small sample sizes. The sensitivity results also highlight a central modeling issue for the intended empirical setting: **when the dataset is conditional on reaching the formal stage, the selection mechanism must be incorporated (or the estimand redefined)** to obtain calibrated credible intervals.

Recommended next step: extend the likelihood to condition on being observed at the formal stage (equivalently, incorporate the probability that an auction reaches the formal stage), then rerun the sensitivity exercises and proceed to covariates and richer bidder heterogeneity.