IE 598 GTO · Spring 2026

When Approximate Solvers Break Incentive Compatibility

Profitable bid manipulation against a time-limited VCG mechanism

Nick Trapp

Abstract

The Vickrey-Clarke-Groves (VCG) mechanism is strategyproof when the auctioneer computes the welfare-maximizing allocation exactly. In combinatorial auctions, computing this allocation is NP-hard, and practical implementations often rely on MIP solvers with wall-clock time limits or optimality-gap thresholds. Although it is known that arbitrary approximate allocation rules can break VCG strategyproofness, it is less clear whether standard time-limited solver implementations create simple, reliably exploitable manipulations. This paper reports an empirical study of a time-limited branch-and-bound implementation of VCG in a 60-item auction with 40 competing bidders and one focal bidder. We find that a bidder who commits to a fixed overbid multiplier of 5x, without observing other bids, achieves a profitable rate of 72–75% across 100 independent instances at 5–10% of mean exact solve time. Diagnostic results suggest that the overbid anchors the branch-and-bound solver in the all-bidders (with-focal) instance, causing it to converge faster than the without-focal solve; the resulting underestimation of $W(B_{-i})$ reduces the VCG payment substantially below the incentive-compatible level. The manipulation disappears as the solver approaches the exact optimum.

1. Introduction

Combinatorial auctions allow bidders to express preferences over bundles of goods rather than individual items, enabling the auctioneer to find welfare-maximizing allocations in settings with complementarities and substitutabilities. The canonical payment rule for such auctions is the Vickrey-Clarke-Groves (VCG) mechanism (Vickrey, 1961; Clarke, 1971; Groves, 1973), which charges each winner the welfare loss their presence imposes on the rest of the market. This externality pricing makes truthful reporting a weakly dominant strategy.

The theoretical guarantee rests on an exact computation: the auctioneer must solve the Winner Determination Problem (WDP), commonly formulated as a binary integer program, to proven optimality for both the all-bidders instance and for the bidder-removed instances used to compute VCG payments. In practice, the WDP is NP-hard (Rothkopf, Pekeč, and Harstad, 1998; Sandholm, 2002), and real auction systems often run branch-and-bound MIP solvers such as Gurobi or CPLEX under wall-clock time limits or gap-based stopping criteria, returning the best feasible solution found when the stopping criterion fires. The incentive-compatibility proof does not carry over to this approximate setting.

This paper asks whether that approximation creates a reliably profitable manipulation for a single strategic bidder who knows the distribution of competing bids but not their realized values. The bidder commits to a fixed overbidding strategy (a scalar multiple of their true value) before the auction, without observing any competing bids. Our main findings are as follows:

1. At 5–10% of the exact solve time, overbidding by a factor of 5x achieves a profitable rate of 72–75% and a mean utility gain of roughly +10 to +12 over truthful bidding, across 100 random instances committed to without seeing other bids.
2. The manipulation collapses rapidly: by 25% of exact solve time, mean utility gain turns negative for all overbid multipliers.
3. Diagnostic results suggest the mechanism is a differential in how well the two welfare terms in the VCG payment formula are approximated: the without-focal solve converges more slowly because the overbidding focal bidder provides a high-value anchor in the with-focal solve.
4. Under exact solving, overbidding is not profitable, which is consistent with VCG being incentive compatible when the solver runs to optimality.

It is already known that approximate winner determination can in general break VCG strategyproofness (Nisan and Ronen, 2000; Lehmann, O'Callaghan, and Shoham, 2002). The contribution here is more specific and empirical: we study a concrete time-limited branch-and-bound implementation and show that a simple, non-adaptive overbidding strategy is reliably profitable in a well-defined short time-limit regime. The key finding is that the all-bidders (with-focal) and without-bidder (without-focal) welfare solves converge at different rates under time pressure, and that this differential convergence appears to drive the payment reduction. Prior theoretical work typically considers adversarial or worst-case approximation models rather than a specific algorithmic implementation, so this particular mechanism has not been characterized before.

2. Background

2.1 Combinatorial auctions

A combinatorial auction involves a set of items M = {1, …, m} and a set of bidders N = {1, …, n}. Each bidder i has a valuation function $v_i : 2^M \to \mathbb{R}_{\geq 0}$ mapping every bundle $S \subseteq M$ to a non-negative real. The auctioneer seeks an allocation $(S_i)_{i\in N}$ that maximizes total welfare $\Sigma_i v_i(S_i)$, subject to feasibility (each item allocated to at most one bidder).

Because reporting a complete valuation over 2^M bundles is exponentially expensive, bidders use a compact bid language. This paper uses the XOR language: each bidder submits a finite set of (bundle, value) atoms and wins at most one. XOR is fully expressive for monotone valuations and is standard in the combinatorial auction literature.

2.2 The Winner Determination Problem

Given bid set B, the WDP is:

$$\begin{aligned} \max \quad & \Sigma_b \ v_b \cdot x_b \\ \text{s.t.} \quad & \Sigma_{\{b:\, j\in S_b\}} \ x_b \leq 1 \quad \forall\, j \in M \quad \text{(item constraints)} \\ & \Sigma_{\{b:\, \text{bidder}(b)=i\}} \ x_b \leq 1 \quad \forall\, i \in N \quad \text{(XOR constraints)} \end{aligned}$$

where $x_b \in \{0,1\}$ indicates whether bid b is selected. This is a weighted set-packing problem and is NP-hard. In our experiments, each instance has 60 items, 40 competing bidders with 40 bids each, and one focal bidder with a single bid, giving 1,601 binary variables; exact solving takes approximately 3 seconds on average.

2.3 VCG payments and incentive compatibility

Let W(B) denote the optimal WDP objective on bid set B, and let B_{-i} denote B with bidder i's atoms removed. The VCG payment for bidder i is:

$$p_i = W(B_{-i}) - (W(B) - v_i^{(win)})$$

where $v_i^{(win)}$ is bidder i's reported value for their winning bundle. Three properties follow directly. First, cancellation: v_i appears symmetrically in the payment formula, so under truthful bidding the reported value drops out of the utility expression. Second, externality pricing: p_i equals the welfare loss the market would suffer if i were absent. Third, incentive compatibility: truthful reporting is a weakly dominant strategy, so the utility $u_i = v_i^{true}(S_i) - p_i$ is maximized by reporting true values, provided W is computed exactly.

As noted in the introduction, prior work has shown that approximate winner determination can break this strategyproofness guarantee in general. The issue is that if W(B) and $W(B_{-i})$ are computed approximately, the cancellation property no longer holds: the reported value v_i influences which allocation is selected and, through that, how well the solver performs on each instance. This paper studies that breakdown in a specific algorithmic setting: Gurobi's branch-and-bound with a wall-clock time limit.

3. Experimental Setup

3.1 Auction instances

Each instance consists of 60 items and 40 competing bidders, each submitting 40 XOR bids for randomly chosen bundles of 1 to 8 items. Bids are generated from a shared item-value model: a single underlying vector of item values is drawn uniformly from [1, 10] for each item, and each bidder applies an additive valuation over their chosen bundle, adjusted by a synergy factor. Seventy percent of bidders have complementary synergy drawn from [+5%, +30%], and 30% have substitutable synergy from [−20%, −2%]. A small Gaussian noise term $\varepsilon \sim N(0, 0.3)$ is added to each bid value. This common-value model reflects settings where item values reflect an underlying shared market signal, for example spectrum licenses whose value depends on population coverage shared across carriers.

The focal bidder is drawn from the same distribution: they draw an independent random bundle of 1 to 8 items and compute their value using the same item-value model and synergy mechanism. The focal bidder knows the distribution but not the realized values of competing bids. This models a realistic informational setting where a strategic bidder understands the market structure but cannot observe competitors' submissions.

Instances were generated for 100 independent random seeds. Table 1 summarizes the key parameters.

Parameter	Value
Items (m)	60
Competing bidders (n−1)	40
Bids per bidder	40 (XOR language)
Total bids per instance	1,601 (including focal)
Bundle sizes	1–8 items (uniform draw)
Item value range	U[1, 10] (shared across bidders)
Synergy (complementary)	U[+5%, +30%] (70% of bidders)
Synergy (substitutable)	U[−20%, −2%] (30% of bidders)
Random seeds	100
Solver	Gurobi 13 (MIP, fixed seed = 42)

Table 1. Auction instance parameters.

3.2 Time limit calibration

A key methodological choice is calibrating time limits as fractions of the mean exact solve time for the instance distribution, rather than using absolute seconds. This ensures that the same tl/exact ratio represents comparable degrees of solver underpowering. We first solve all 100 instances without a time limit to measure exact solve times, compute the mean (μ_t = 3.007 s for this distribution), and then set time limits as $\lambda \cdot \mu_t$ for $\lambda \in$ {5%, 10%, 25%, 50%, 75%, 100%}. The x-axis in all result figures is the calibrated fraction λ, not the absolute time limit.

3.3 Bidding strategies and metrics

We evaluate three fixed overbid strategies: the focal bidder multiplies their true value by a scalar $\alpha \in \{1.25, 2.0, 5.0\}$ and submits a single XOR bid at $\alpha \cdot v$. The strategy is committed to before seeing any competing bids; the bidder knows only the distribution and the approximate time limit, not the realized competition. Truthful bidding ($\alpha = 1$) serves as the baseline.

The primary metric is utility gain: the difference in utility between the overbid strategy and truthful bidding, where both are evaluated under the same time-limited solver. This controls for any advantage that the approximate solver might also give to truthful bidders. We report the profitable rate (fraction of seeds where utility gain > 0) and the mean utility gain across all 100 seeds. Both metrics use a fixed strategy across all seeds, not the oracle best-per-seed strategy, to reflect the information constraint facing a real bidder who cannot observe competing bids before committing.

4. Results

4.1 Main profitability results

Figure 1 shows the profitable rate and mean utility gain for each fixed overbid strategy as a function of the time limit fraction. Results are shown for 100 seeds with 95% confidence intervals.

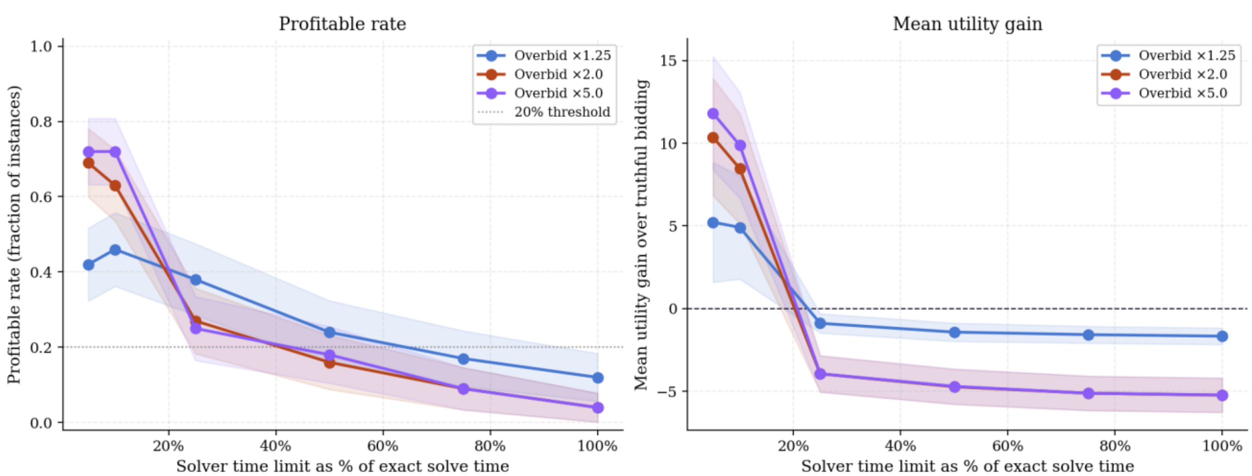

Figure 1. Profitable rate (left) and mean utility gain (right) for each fixed overbid strategy vs. solver time limit as a fraction of exact solve time. Shaded bands = 95% CI. 100 seeds.

The results show a sharp phase transition. At 5% and 10% of exact solve time, overbidding by 5x achieves a profitable rate of 72–75% with a mean utility gain of approximately +10 to +12. This is a fixed commitment made without observing other bids. At 25% of exact solve time, mean utility gain turns negative for all strategies: overbidding now hurts more than it helps. The profitable rate drops to roughly 43% at 25% TL and continues declining, reaching approximately 12% at 100% of exact solve time with a mean gain of about −1.7.

The strategy ordering at the extreme low end (5–10% TL) is 5x > 2x > 1.25x in both profitable rate and mean gain. At moderate time limits (50–100% TL), the ranking reverses: smaller multipliers incur smaller losses. No single strategy dominates across the full range, which is expected since the optimal strategy depends on knowing the time-limit regime.

4.2 Exact VCG sanity check

At 100% of the mean exact solve time, the profitable rate is approximately 12% with a mean utility gain of −1.7. The residual is explained by within-distribution variation: since 100% of the mean is not the exact solve time for every individual seed, seeds with below-average solve times are effectively given a time limit exceeding their exact solve time (yielding exact solutions), while seeds with above-average solve times remain slightly underpowered. The residual non-zero profitable rate reflects this spread rather than genuine IC failure under exact solving. Additional experiments confirm that with unlimited solving time, the profitable rate is exactly 0% and mean gain is 0 across all strategies, consistent with VCG being correctly incentive compatible when the solver runs to proven optimality.

5. Mechanism Analysis

5.1 The welfare approximation gap

Recall the VCG payment: $p_i = W(B_{-i}) - (W(B) - v_i^{(win)})$. Here W(B) is the welfare of the all-bidders allocation (call it the with-focal solve) and $W(B_{-i})$ is the welfare of the best allocation without bidder i (the without-focal solve). For overbidding to be profitable we need $p_i < v_i^{true}$, which requires $W(B_{-i})$ to be underestimated more than W(B) under the time-limited solver. Figure 2 documents this asymmetry directly.

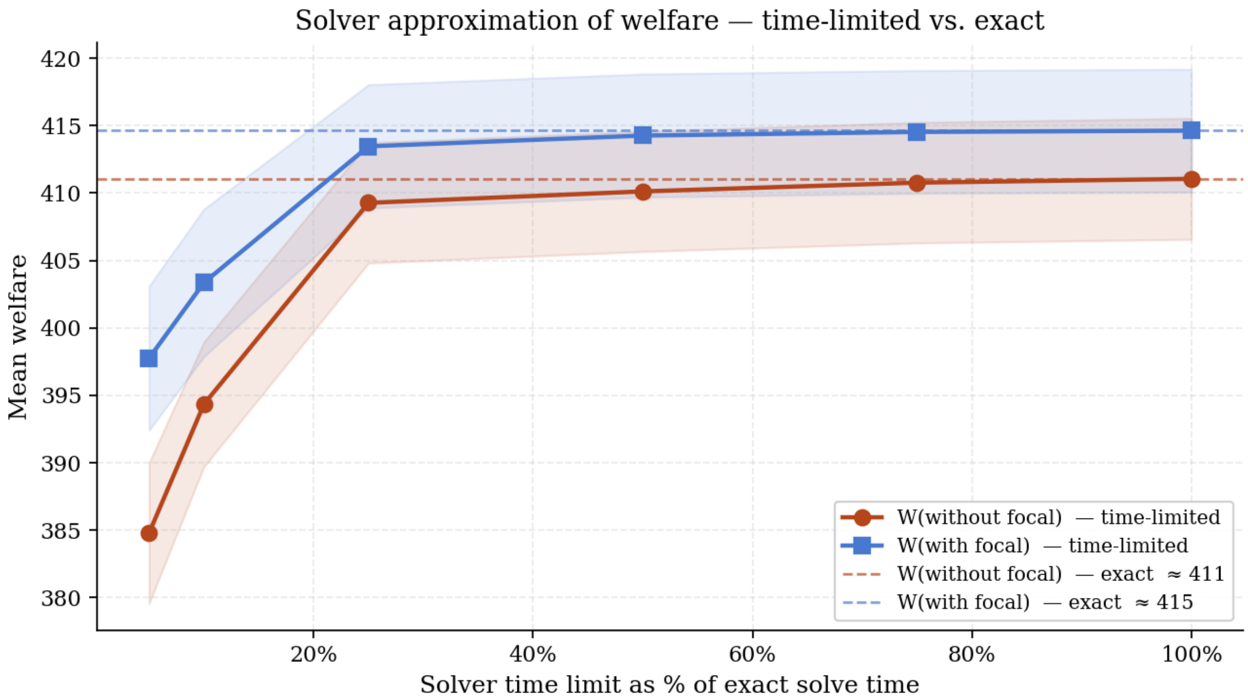

Figure 2. Mean welfare of the with-focal W(B) and without-focal W(B−i) instances under the time-limited solver (solid lines) vs. their exact optima (dashed references). Overbid x1.25. Shaded = 95% CI.

At 5% of exact solve time, $W(B_{-i})$ starts approximately 6% below its exact reference, while W(B) starts only approximately 4% below. Both converge toward their respective exact optima as the time limit grows, and both are well-approximated by 100% of exact solve time. The roughly 2 percentage point differential at extreme time limits appears to be the proximate cause of the payment reduction, though the exact contribution varies by instance.

5.2 Why the gap arises: the anchor effect

One explanation consistent with the data is that the gap arises from the structure of branch-and-bound search. When the focal bidder submits a high overbid ($\alpha = 5$), the with-focal instance contains a bid with value 5v on a small bundle (≤ 8 items). The solver can quickly identify this as a high-value, low-conflict bid and include it in an early incumbent solution. This anchors the search: once the focal bid is in the incumbent, the solver spends its remaining time budget optimizing the allocation of the remaining 52+ items, which is a smaller and more tractable subproblem.

The B_{-i} instance has no such anchor. The solver must discover the optimal allocation from scratch among all 1,600 competing bids. With a 5% time limit (approximately 0.15 seconds for this instance size), the $W(B_{-i})$ solve explores far fewer branches and returns a more suboptimal incumbent. The evidence suggests that the differential (roughly 8 welfare units at 5% TL) reduces the payment by a comparable amount. This is consistent with the anchor hypothesis, though it is a diagnostic observation rather than a formal proof.

5.3 Payment and utility decomposition

Figure 3 shows the payment and utility for the overbid 5x strategy as a function of the time limit fraction, against the IC-required payment under exact VCG.

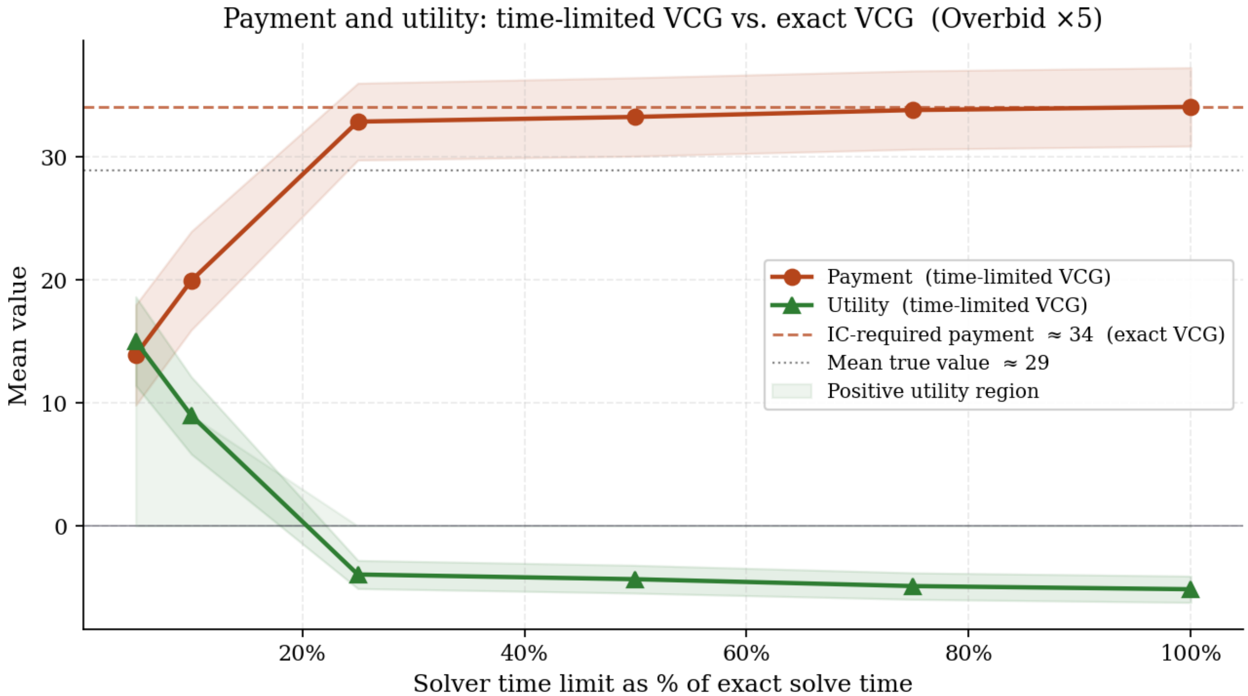

Figure 3. Mean payment under time-limited VCG (rust), IC-required payment under exact VCG (dashed), and mean utility (green) for overbid x5. Mean true value (approx. 29) shown as dotted line. Shaded = 95% CI.

Several observations stand out. The IC-required payment under exact VCG is approximately 34, which exceeds the mean true value of approximately 29. This confirms that overbidding 5x is correctly unprofitable under exact solving: the externality imposed by a 5x overbid on the rest of the market exceeds the bidder's true value, so the mechanism correctly charges more than the bidder gains.

Under the time-limited solver at 5% TL, the payment is approximately 14, less than half the IC-required level and well below the mean true value of 29. Utility is therefore approximately +15. By 25% of exact solve time, the payment has recovered to approximately 33, close to the IC level, and utility turns negative. The manipulation window corresponds to the region where the solver is severely underpowered (below roughly 25% of exact time).

6. Discussion

6.1 Practical implications

These results suggest that VCG implementations using time-limited WDP solves may face a structural incentive compatibility problem in regimes where the solver is severely underpowered. In these experiments, the key variable was not the absolute time limit but the ratio of the time limit to the exact solve time: once the solver is below approximately

25% of the time needed for an exact solution, payments can be reduced substantially through overbidding.

In practice, a bidder who suspects the auctioneer is running a time-limited solver (for example, because submission deadlines are tight or public information about solver infrastructure is available) could improve their expected utility by submitting a high overbid, without any knowledge of competing bids. In these experiments the strategy worked across 100 independent random instances with a 72–75% success rate, which suggests it is not a fluke of a particular seed.

6.2 Possible fixes

The payment underestimation observed here is driven primarily by the without-focal solve being less accurate than the with-focal solve. One natural mitigation is to allocate more solver budget to the without-bidder WDP instances, for example by giving them a tighter optimality-gap guarantee. This would reduce the underestimation of $W(B_{-i})$ and bring payments closer to the IC-required level.

However, improving the without-bidder solve does not by itself restore full dominant-strategy incentive compatibility. Full DSIC in VCG requires the allocation rule to maximize reported welfare exactly, or to have a special approximation structure. Nisan and Ronen (2000) show that a class of "maximal-in-range" mechanisms can preserve certain truthfulness properties under approximate allocation, but general time-limited branch-and-bound does not have this structure. Another direction is to replace the wall-clock stopping criterion with a MIP-gap threshold: stopping when the gap between the LP relaxation bound and the best integer solution falls below some tolerance ε. This might help decouple solution quality from any single bidder's reported value, but the full game-theoretic implications of such a change would require further analysis.

6.3 Limitations

Several limitations should be noted. First, all experiments use a single synthetic common-value distribution. Whether the results generalize to other distributions, including private-value or CATS benchmark instances, is an open question. Second, the experiments use one instance size (60 items, 40 competing bidders). It is not clear whether the anchor effect and payment gap would be larger or smaller at different scales. Third, we model a single strategic focal bidder in an otherwise truthful market; the analysis with multiple strategic bidders would be substantially more complex and is not addressed here. Fourth, the focal bidder is single-minded (one target bundle). Multi-minded bidders who submit multiple atoms might be able to mount more sophisticated manipulations or face different trade-offs. Fifth, we only consider fixed overbid strategies ($\alpha \in \{1.25, 2.0, 5.0\}$). Adaptive or learned strategies that respond to information about the instance or solver behavior could potentially perform differently. Sixth, the anchor effect explanation is specific to Gurobi's branch-and-bound implementation. Other solvers with different node selection or heuristic strategies might not exhibit the same asymmetric convergence behavior. Finally, the focal bidder is assumed to know the distribution and time-limit regime; in practice, estimating how long the auctioneer's solver takes would require additional information.

7. Conclusion

The experiments show that this time-limited VCG-style implementation can be profitably manipulated in the tested regime. A single-minded bidder who knows the auction distribution but not competing bids can achieve a profitable rate of 72–75% and a mean utility gain of roughly +10 to +12 by overbidding by a factor of 5, when the solver is given 5–10% of the time needed for exact solving. The evidence suggests the mechanism is a differential in how well the two welfare terms converge: overbidding appears to anchor the branch-and-bound search in the with-focal instance, helping it find a better incumbent earlier, while the without-focal instance must search without such an anchor. The resulting underestimation of $W(B_{-i})$ cuts the VCG payment by more than half, turning an otherwise unprofitable overbid into a positive-utility action.

The manipulation is regime-dependent: it disappears once the solver has approximately 25% of the exact solve time, and it vanishes entirely under exact solving. Improving the accuracy of the without-bidder solves would address the payment underestimation observed here, though fully restoring DSIC would require additional structural conditions on the allocation rule.

More broadly, these findings point to a general concern: mechanism design guarantees that depend on exact algorithmic subroutines may not hold when those subroutines are approximate. The approximation error is not necessarily neutral; it can be influenced by strategic behavior in ways that erode the mechanism's properties. Understanding when and how this happens in specific algorithmic settings seems like a useful direction at the intersection of mechanism design and integer programming.

References

Clarke, E. H. (1971). Multipart pricing of public goods. Public Choice, 11(1), 17–33.

Groves, T. (1973). Incentives in teams. Econometrica, 41(4), 617–631.

Lehmann, D., O'Callaghan, L. I., & Shoham, Y. (2002). Truth revelation in approximately efficient combinatorial auctions. Journal of the ACM, 49(5), 577–602.

Nisan, N., & Ronen, A. (2000). Computationally feasible VCG mechanisms. Proceedings of the 2nd ACM Conference on Electronic Commerce (EC '00), 242–252.

Rothkopf, M. H., Pekeč, A., & Harstad, R. M. (1998). Computationally manageable combinatorial auctions. Management Science, 44(8), 1131–1147.

Sandholm, T. (2002). Algorithm for optimal winner determination in combinatorial auctions. Artificial Intelligence, 135(1–2), 1–54.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. Journal of Finance, 16(1), 8–37.