Certify What the Budget Can Reach: Peer Assessment in Federated Data Markets

Anonymous authors
Paper under double-blind review

ABSTRACT

In federated data markets, evaluating decentralized client contributions without central ground truth or public validation sets is a core challenge. Peer prediction addresses this by rewarding informative reports via peer cross-validation, yet its theoretical guarantees often become operationally unusable in practice. Because distributed supplier pairs share only scarce task overlap, the platform must ration a limited assessment budget across many candidate pairs. Existing mechanisms assume fixed task assignments, while sequential allocation methods merely reduce generic uncertainty; neither assesses whether the remaining budget and local overlap can actually complete statistically defensible incentive evidence. Consequently, cautious schedulers often exhaust the entire budget without certifying any supplier, even when multiple certificates were affordable. We formalize this missing objective as budget-reachable certification under adaptive sampling and heterogeneous finite pools. We propose BRCert, which pairs simultaneous finite-population confidence bounds with a lookahead estimator of each edge's remaining certification cost. BRCert prunes unreachable pairs and concentrates purchases on the most affordable reachable certificates without modifying the underlying payment rule. Across five real-world markets, BRCert achieves the highest mean certified coverage with zero false certificates under exact audits, outperforming the strongest baseline by 1.6 percentage points on average in the main setting. The gain increases to 1.9 percentage points with 50 candidate edges, demonstrating that finite-batch incentive design must prioritize completed evidence over diffuse exploration.

1 INTRODUCTION

Markets for labels, forecasts, sensor readings, and federated model updates often lack an immediately verifiable ground truth. Peer-prediction mechanisms address this obstacle by comparing reports from agents who observe related tasks and correcting raw agreement for uninformative coordination (Dasgupta & Ghosh, 2013; Shnayder et al., 2016; Kong & Schoenebeck, 2019). This principle has recently been instantiated in federated learning without a public validation set or known report distribution (Witt et al., 2026).

However, a truthful payment rule does not by itself produce usable evidence that truthfulness was rewarded. Practical markets repeatedly evaluate heterogeneous suppliers, each pair shares only a finite set of tasks, and every revealed comparison consumes budget and overlap (Chen et al., 2020; Feng et al., 2022). If purchases are spread across all uncertain pairs, every confidence interval may remain inconclusive. If they are concentrated greedily on the largest estimated margin, the platform may exhaust its budget on a pair whose certificate cannot be finished. Both failures have the same costly consequence: payments and scarce shared tasks are consumed, yet the platform cannot defend any supplier evaluation. This can misclassify both high-scoring and underexplored suppliers.

This operational failure falls between existing research programs. Peer-prediction theory characterizes expected incentives under a fixed assignment or report distribution, but generally leaves the

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.

evidence-acquisition schedule exogenous (Kong & Schoenebeck, 2019; Chen et al., 2020; Feng et al., 2022). Finite-sample estimators and confidence sequences quantify uncertainty for a given stream, but do not determine which of many exhaustible pair streams should receive the next sample (Howard et al., 2021; Waudby-Smith & Ramdas, 2024). Bandit and thresholding methods allocate observations adaptively, yet their usual objectives, such as reward, uncertainty reduction, or asymptotic identification, do not value the completion of a valid certificate before a hard terminal budget (Alieva et al., 2021; Chen et al., 2025; Russo et al., 2025). Consequently, none of these tools distinguishes a promising edge from a certificate that is still affordable.

Closing this gap is challenging for three coupled reasons. First, there is no verified label with which to audit reports, so the target must be an incentive margin induced by the peer-payment rule itself (Chen et al., 2020). Second, pair pools are heterogeneous and finite: additional evidence has edge-specific cost and may simply not exist (Waudby-Smith & Ramdas, 2024). Third, allocation is adaptive, so concentrating samples based on observed reports must not invalidate the final statistical claim (Howard et al., 2021). A useful scheduler must therefore reason simultaneously about payoff separation, finite-population uncertainty, local pool exhaustion, and a shared global budget.

To address this gap, we study *adaptive incentive certification*: how should a platform allocate reference peers and shared tasks to maximize the number of evaluations for which truthful reporting has a statistically supported payoff advantage? We retain a fixed multi-task peer-prediction rule, Knowledge-Free Correlated Agreement (KFCA) in our main instantiation (Witt et al., 2026), and learn only its evaluation schedule. Certification requires a simultaneous lower bound on the worst-case truthful payoff advantage to exceed a target ε.

Our key insight is that this is a *reachability* problem before it is an exploration problem. Instead of asking only which edge is promising or uncertain, the scheduler must determine which valid certificate can still be completed before its local pool or the global budget expires. The payment formula, scale, and current assignment remain fixed before reports arrive; adaptation changes only where future evidence is purchased.

Our contributions are summarized as follows:

- We formulate *budget-reachable incentive certification*, making completed statistical evidence, rather than score, reward, or generic uncertainty reduction, the allocation objective in repeated peer evaluation. We construct pair-specific certificates for randomly ordered finite overlap pools and prove their simultaneous validity under adaptive edge selection.
- We propose BRCERT, which converts empirical incentive margins and finite-population confidence radii into the additional evidence required for certification. It removes edges that cannot be completed within their remaining overlap or the global budget, then prioritizes the most affordable reachable certificate.
- We evaluate nine schedulers on five real annotation markets, including CIFAR-10N as a sparse large-alphabet stress test. BRCERT obtains the highest mean certified-edge coverage on all five datasets and produces no false certificate under exact finite-pool audits. It outperforms the strongest baseline by 1.6 percentage points on average in the main setting, and the gain increases to 1.9 percentage points with 50 candidate edges.

2 RELATED WORK

Peer prediction and correlated agreement. Peer prediction originated as a way to elicit subjective information without verification (Prelec, 2004; Miller et al., 2005; Jurca & Faltings, 2009; Witkowski & Parkes, 2012). Multi-task mechanisms distinguish informative agreement from agreement caused by report marginals (Dasgupta & Ghosh, 2013; Shnayder et al., 2016), while later work studies robust and detail-free variants for heterogeneous reporting environments (Radanovic & Faltings, 2015; Kong & Schoenebeck, 2019; Liu & Chen, 2018). Empirical studies also show that incentive magnitude and payment variance can differ substantially across peer mechanisms (Gao et al., 2016; Wang et al., 2021). KFCA fixes an equality-based score and obtains strong truthfulness under a categorical-world sign condition and an honest-majority argument (Witt et al., 2026). We use KFCA as a modular payment rule; our object of study is which heterogeneous pairs should supply the shared reports and how much evidence should be purchased.

Recent finite-sample analysis characterizes which mutual-information quantities admit unbiased estimators, including variance-minimal and random-stopping constructions (Aznag et al., 2026). Those results concern a fixed joint report distribution. We instead allocate samples across many finite pair pools and certify a decision-relevant payoff gap. Classical and modern without-replacement concentration bounds provide the statistical basis for our finite-pool certificate (Serfling, 1974; Bardenet & Maillard, 2015); tighter betting and confidence-sequence methods are complementary extensions (Howard et al., 2021; Waudby-Smith & Ramdas, 2024).

Data markets and learning incentives. Peer prediction has been applied directly to one-shot and repeated acquisition of datasets without validation data (Chen et al., 2020). This differs from contribution-based data markets that price marginal predictive value or approximate Shapley values (Agarwal et al., 2019; Ghorbani & Zou, 2019; Jia et al., 2019; Pei, 2022). Sequential mechanisms have learned posted rewards or payment intensity (Liu & Chen, 2017; Hu et al., 2018), and recent work studies how learning agents behave inside the named Correlated Agreement mechanism (Feng et al., 2022). BRCERT leaves the payment rule unchanged and decides only where a finite assessment budget can complete a valid certificate.

Sequential allocation. Modern bandit methods distinguish cumulative reward maximization from fixed-budget pure exploration (Lattimore & Szepesvári, 2020; Alieva et al., 2021). Our target is closer to thresholding than reward maximization: APT is a canonical named rule (Locatelli et al., 2016), while recent Confident Sample work studies adaptive threshold sampling under structured actions (Chen et al., 2025). Selecting a peer edge also creates overlap and graph constraints. Recent methods such as TaS-FG explicitly study pure exploration with feedback graphs (Russo et al., 2025), and new theory quantifies when adaptivity can improve identification complexity (Maiti et al., 2026). These settings do not include finite certification pools or peer payments; they motivate the closest allocation baselines rather than providing drop-in mechanisms. Unlike asymptotic identification, BRCERT explicitly accounts for a hard terminal budget and an exhaustible sample pool.

Sequential incentives and federated learning. Truth inference and label aggregation estimate latent answers or worker reliability from redundant annotations (Dawid & Skene, 1979; Raykar et al., 2010; Zheng et al., 2017; Li et al., 2022); certification instead asks whether a fixed payment favors truthful reports. Federated learning remains a possible application because contributions are heterogeneous and hard to validate (McMahan et al., 2017; Kairouz et al., 2021; Sim et al., 2020; Zhan et al., 2022; Wu et al., 2024), but our experiments use real annotation logs and require no model training.

3 REPEATED DATA MARKET

3.1 AGENTS, TASKS, AND REPORTS

There are N persistent data suppliers indexed by $i \in [N]$ and decision epochs $t = 1, \dots, T$. A task k has an unobserved categorical state $Y^k \in [L]$. When supplier i is assigned task k, it observes a private signal $Z_i^k \in [L]$ after exerting the required effort and submits report $R_i^k \in [L]$. A reference edge $e = (i, j)$ is evaluated on tasks observed by both endpoints. The platform does not observe Y^k.

For each feasible edge e, the market contains a finite pool $\mathcal{P}_e$ of M_e shared-task tuples. Pair pools may differ in size and report structure. Let $c_e > 0$ be the cost of revealing one complete evaluation tuple for edge e. At epoch t, the platform chooses a graph $G_t \subseteq E$ and integer batch sizes $b_{e,t}$ before observing any report in that epoch, subject to

$$\sum_{e \in G_t} c_e b_{e,t} \le B_t, \qquad \deg_{G_t}(i) \le d_i, \tag{1}$$

where B_t is the evaluation budget and d_i limits supplier workload.

Assumption 1 (Random-order finite pools)**.** *For every edge, the order of tuples in $\mathcal{P}_e$ is sampled uniformly without replacement and independently of the scheduler. The scheduler may adaptively choose the next edge using past revealed tuples, but cannot choose which unseen tuple of that edge is revealed next.*

The target is the mean incentive margin of the declared finite batch, not an extrapolation to future tasks. This is the natural contract when a marketplace certifies a fixed uploaded batch or a logged annotation pool. Chronological deployment requires a separate drift model and is not covered by this guarantee.

3.2 Fixed peer payment

Our main instantiation uses the multi-task construction underlying KFCA (Witt et al., 2026). A sample contains a shared bonus task k and independently selected penalty tasks p and q. For reports from suppliers i and j, the unscaled score is

$$X_e = ind\{R_i^k = R_j^k\} - \mathbb{I}\{R_i^p = R_j^q\}, \qquad X_e \in [-1, 1]. \tag{2}$$

Payments are $a_e X_e + h_e$, where $a_e > 0$ and the base payment h_e are posted before reports. Our allocation method never changes this formula in response to current reports. The categorical-world condition of KFCA remains an applicability condition; adaptive scheduling cannot manufacture truthful incentives for a pair whose signal structure violates the payment rule's assumptions.

Let $f : [L] \to [L]$ denote a deterministic reporting deviation by the evaluated supplier i, with the reference supplier truthful. Applying f to the observed i-side signal in Eq. (2) gives the counterfactual score $X_e(f)$. For a finite declared deviation class $\mathcal{F}$ containing the identity map $f^\star$, define

$$\Gamma_e(f) = \frac{a_e}{M_e} \sum_{s \in \mathcal{P}_e} [X_{e,s}(f^\star) - X_{e,s}(f)], \qquad \Gamma_e = \min_{f \in \mathcal{F} \setminus \{f^\star\}} \Gamma_e(f). \tag{3}$$

Γ_e is the worst-case expected reporting advantage within $\mathcal{F}$. For small categorical alphabets we enumerate all deterministic maps; for larger alphabets we use a declared attack family including constant, permutation, coarsening, and label-flip strategies.

Scope of the margin. The margin compares reporting strategies under the prescribed effort and reference behavior. It is not a universal certificate of semantic correctness, nor does it identify ground truth. Effort incentives additionally require the margin in expected payment to exceed the cost saved by shirking. Coalition guarantees require conditions on the local peer graph and are addressed separately in Section 6.

4 Finite-Sample Incentive Certification

For $n_e(t)$ completed samples on edge e, let

$$\widehat{\Gamma}_e(f; t) = \frac{a_e}{n_e(t)} \sum_{s=1}^{n_e(t)} [X_{e,s}(f^\star) - X_{e,s}(f)], \quad \widehat{\Gamma}_e(t) = \min_{f \neq f^\star} \widehat{\Gamma}_e(f; t). \tag{4}$$

The counterfactual terms are available because every $f \in \mathcal{F}$ is applied to already observed categorical reports; no ground-truth label is required.

Since $a_e[X_e(f^\star) - X_e(f)] \in [-2a_e, 2a_e]$, a simultaneous Hoeffding–Serfling radius for sampling without replacement is

$$r_e(n, \delta) = 4a_e \sqrt{\frac{M_e - n}{M_e - 1}} \sqrt{\frac{1}{2n} \log\left(\frac{2|E|(|\mathcal{F}| - 1)M_{\max}}{\delta}\right)}. \tag{5}$$

We define

$$\mathrm{LCB}_e(t) = \widehat{\Gamma}_e(t) - r_e(n_e(t), \delta), \qquad \mathrm{UCB}_e(t) = \widehat{\Gamma}_e(t) + r_e(n_e(t), \delta). \tag{6}$$

Definition 1 (Incentive certificate)**.** *For target margin $\varepsilon \geq 0$, edge e is (ε, δ)-certified at time t when $\mathrm{LCB}_e(t) \geq \varepsilon$. It is rejected for this target when $\mathrm{UCB}_e(t) < \varepsilon$. Otherwise it is unresolved.*

Theorem 1 (Simultaneous finite-pool validity)**.** *Under Assumption 1, with probability at least $1 - \delta$, simultaneously for every edge, deviation, and sample count $1 \leq n \leq M_e$ reached by an adaptive scheduler, $|\widehat{\Gamma}_e(f; t) - \Gamma_e(f)| \leq r_e(n_e(t), \delta)$. Consequently, every certified edge satisfies $\Gamma_e \geq \varepsilon$ and every rejected edge satisfies $\Gamma_e < \varepsilon$ on the same event.*

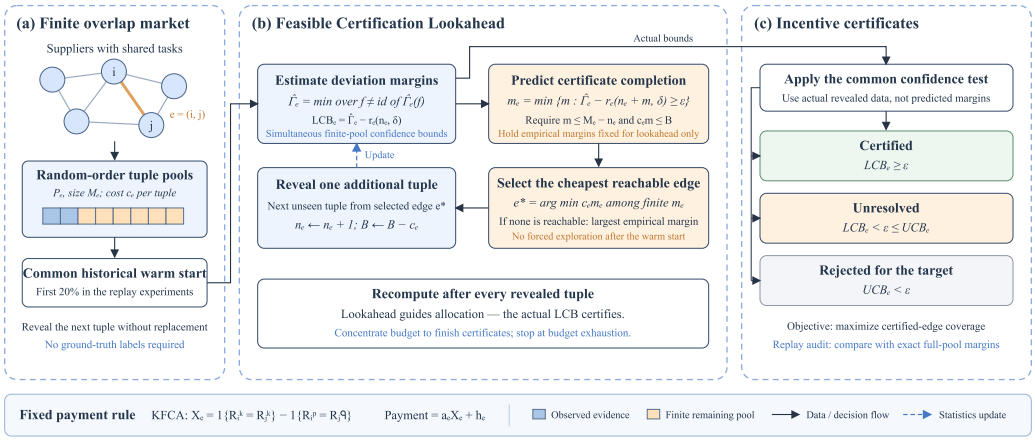


Figure 1: BRCERT overview.

The proof applies the without-replacement concentration bound at each (e, f, n) and unions over all declared finite pools. Adaptive edge choice is harmless because the event already holds for every sample count. The finite-population correction tends to zero only because the estimand is the exact finite pool.

For $\Delta_e = |\Gamma_e - \varepsilon| > 0$, define $m_e^\star$ as the smallest additional sample count for which the current plug-in margin minus $r_e(n_e + m, \delta)$ reaches ε. This quantity is finite only when the edge can be completed within both its remaining pool and the remaining global budget. It is precisely this finite-horizon distinction that motivates our scheduler.

5 BUDGETED ADAPTIVE PEER ASSIGNMENT

5.1 BUDGET-REACHABLE CERTIFICATION

At decision step t, let n_e be the revealed sample count, $R_e = M_e - n_e$ the remaining pool, and B_t the remaining budget. BRCERT holds the empirical deviation margins fixed only for scheduling and computes

$$m_e(t) = \min \left\{ m \in \{0, \dots, \min(R_e, B_t)\} : \min_{f \neq f^\star} \widehat{\Gamma}_e(f; t) - r_e(n_e + m, \delta) \geq \varepsilon \right\}. \tag{7}$$

If the set is empty, $m_e(t) = \infty$. This plug-in calculation is not itself a certificate; it only predicts the cost of reaching the common valid certificate. BRCERT selects the unresolved edge with the smallest finite $c_e m_e(t)$, breaking ties by larger empirical margin. If no edge is currently reachable, it selects the largest empirical margin so that evidence is concentrated on the most promising candidate. Reachability is recomputed after every revealed tuple.

This rule deliberately avoids forced exploration. The initial history split supplies a common warm start, and under a one-shot finite-batch contract, spending scarce budget on every edge may prevent any certificate. BRCERT is therefore designed to maximize usable certificates in the current batch, not to learn all edge means asymptotically.

Figure 1 summarizes the complete workflow. The platform first constructs random-order finite overlap pools, then uses budget reachability only to schedule the next tuple. The final certified, unresolved, or rejected status always comes from the common confidence test applied to actually revealed evidence; predicted margins never replace the certificate.

Proposition 1 (Certificate completion under stable margins)**.** *Condition on a step at which the empirical margin of a reachable edge remains unchanged for its next $m_e(t)$ observations. If* BRCERT *keeps selecting that edge, it is certified within $m_e(t)$ additional observations. On the event of Theorem 1, the completed certificate is valid regardless of the plug-in prediction.*

Algorithm 1 BRCERT: Budget-Reachable Certification

Require: Historical edge statistics, budget B_t, threshold ε, confidence δ
1: Compute empirical margins and valid lower bounds for every edge
2: For each unresolved edge, compute $m_e(t)$ from Eq. (7)
3: **if** at least one edge has finite $m_e(t)$ **then**
4: Select the edge with minimum predicted cost $c_e m_e(t)$
5: **else**
6: Select the available edge with the largest empirical margin
7: **end if**
8: Reveal its next random-order tuple and update statistics
9: Stop at budget exhaustion; return all edges whose valid LCB exceeds ε

The proposition separates scheduling from validity: inaccurate reachability estimates can reduce coverage, but cannot create a false certificate. BRCERT makes no asymptotic best-arm claim; its purpose is to finish feasible certificates under a hard finite horizon.

6 STRATEGIC SCOPE IN A REPEATED MARKET

6.1 STAGE-WISE INCENTIVES

The schedule for epoch t is measurable with respect to history H_{t-1} and frozen before new reports arrive. Conditional on H_{t-1}, current matching and sample counts are therefore unaffected by a current report.

Corollary 1 (Certified stage-wise margin)**.** *On the event of Theorem 1, if edge $e = (i, j)$ is (ε, δ)-certified and the reference behavior and signal model remain within the certification assumptions, truthful reporting by i yields at least ε more expected current-batch payment than any $f \in \mathcal{F} \setminus \{f^\star\}$.*

The result is deliberately local. It does not assert that a high score means high semantic data quality, that arbitrary coalitions cannot coordinate, or that a forward-looking supplier cannot trade current payment for a more favorable future schedule.

6.2 FORWARD-LOOKING MANIPULATION

An adaptive scheduler creates a new channel absent from a one-shot mechanism: current reports update the state used for future peer and sample allocation. Let $V_i(H_t)$ be supplier i's discounted continuation utility after epoch t, and define the maximum scheduling influence of one reporting deviation by

$$D_{i,t} = \sup_{f \in \mathcal{F}} \left| \mathbb{E}[V_i(H_t^f) - V_i(H_t^{f^\star}) \mid H_{t-1}] \right|. \tag{8}$$

If a certified current margin exceeds $D_{i,t}$, truthfulness remains optimal for the current-plus-continuation utility. Estimating a useful non-vacuous upper bound on $D_{i,t}$ is separate from finite-sample certification. Our finite-batch experiments do not estimate this continuation value and therefore do not claim dynamic incentive compatibility. Establishing such a result for arbitrary patient coalitions would require substantially stronger assumptions and is outside our scope.

6.3 COALITIONS AND HONEST REFERENCES

KFCA's honest-majority argument does not automatically transfer to an adaptively concentrated subgraph. BRCERT certifies the declared unilateral deviation class against the observed reference behavior; it is not coalition-proof. Diversity constraints or random anchors can be added in deployment, but they are outside the present minimal mechanism.

Table 1: Certified-edge coverage (%) with 30 candidate edges, budget 200, and pool cap 30.

Dataset	Uniform	CA-greedy	APT	TaS-FG	Margin	ForcedEX	BRCert
CIFAR-10N	0.0 ± 0.0	19.3 ± 2.4	0.7 ± 0.7	18.0 ± 0.8	24.0 ± 1.2	26.0 ± 0.7	$\mathbf{28.0 \pm 0.8}$
Fact	0.0 ± 0.0	17.3 ± 1.9	6.7 ± 1.8	11.3 ± 0.8	20.0 ± 1.5	18.7 ± 0.8	$\mathbf{22.0 \pm 0.8}$
NetEase-126	0.0 ± 0.0	14.0 ± 1.6	10.0 ± 1.5	9.3 ± 1.2	16.7 ± 2.4	14.0 ± 2.4	$\mathbf{17.3 \pm 1.9}$
NetEase-69	0.0 ± 0.0	15.3 ± 1.7	11.3 ± 1.3	8.0 ± 2.0	15.3 ± 3.1	12.7 ± 3.2	$\mathbf{16.0 \pm 3.2}$
RTE	0.0 ± 0.0	12.0 ± 1.3	8.7 ± 1.3	8.7 ± 0.8	14.0 ± 2.4	12.7 ± 1.9	$\mathbf{14.7 \pm 2.3}$

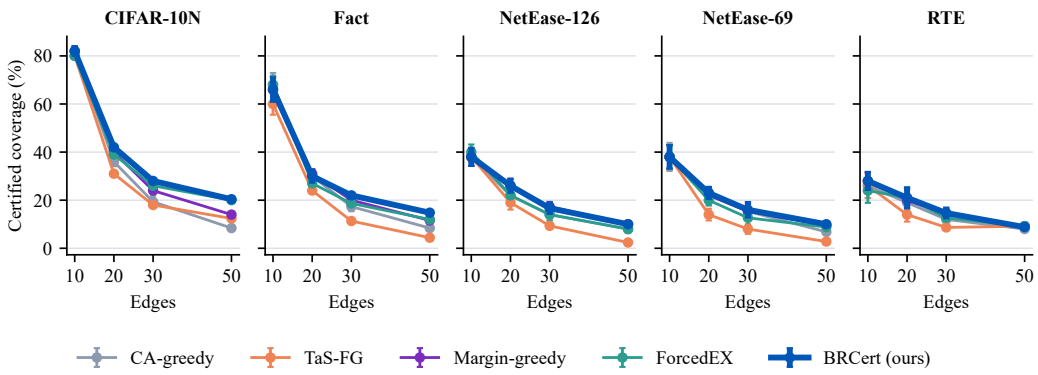


Figure 2: Certified-edge coverage as the candidate set grows.

7 Experimental Design

7.1 Datasets and replay protocol

We replay five real annotation markets without training a predictive model:

- **Fact** contains 214,915 reports from 57 workers on 42,624 knowledge-graph fact-verification items released for the TREC 2013 crowdsourcing setting (Kazai et al., 2013).
- **RTE** contains 8,000 reports from 164 workers on 800 textual-entailment items and is a standard example of non-expert NLP annotation (Snow et al., 2008).
- **NetEase-69** contains 30,353 reports from 327 workers on 4,942 tasks, and **NetEase-126** contains 146,177 reports from 237 workers on 23,851 tasks. Both are capability-specific slices of the large, timestamped NetEaseCrowd collection (Wang et al., 2024).
- **CIFAR-10N** contributes 150,000 human reports from 747 annotators over 50,000 images (Wei et al., 2022). We treat it as an annotation market: image pixels are not loaded and no CIFAR model is trained.

Replay construction. For each seed, we select candidate edges by overlap, randomly permute each shared-task pool, and reserve the first 20% as common history. At most $M \in \{15, 30, 45\}$ complete KFCA tuples are retained per edge. The main setting has 30 candidate edges, $M = 30$, budget 200, $\delta = 0.1$, and $\varepsilon = 0$. We enumerate all deterministic deviations for binary data and constant/permutation deviations for ten-class data. The exact full-pool margin is revealed only after replay to audit certificates.

7.2 Baselines and metrics

All methods share the same history, payment statistic and confidence bound:

- **Margin-greedy** selects the largest empirical incentive margin but ignores whether its certificate is reachable within the remaining horizon. **Uniform** randomly spreads purchases across unresolved edges, the standard non-adaptive control.

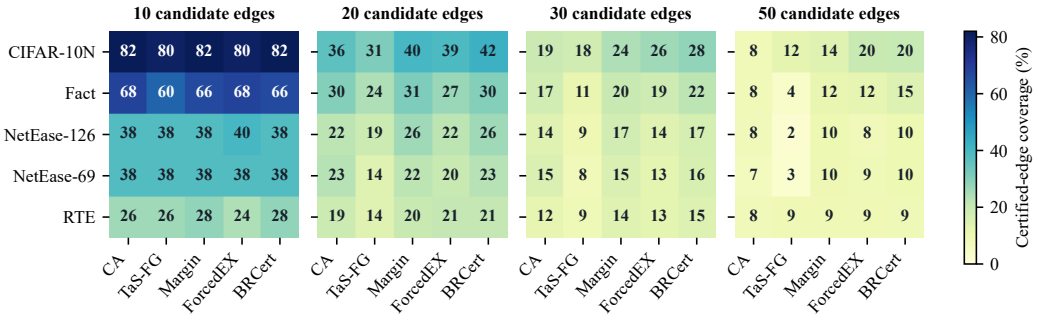


Figure 3: Coverage (%) for five principal schedulers.

Table 2: Budget sensitivity.

Dataset	$B = 100$			$B = 200$			$B = 300$		
	Margin	ForcedEX	BRCert	Margin	ForcedEX	BRCert	Margin	ForcedEX	BRCert
CIFAR-10N	10.0	**13.3**	**13.3**	24.0	26.0	**28.0**	40.0	40.0	**43.3**
Fact	8.0	10.0	**10.7**	20.0	18.7	**22.0**	31.3	30.0	**32.0**
NetEase-126	5.3	4.7	**6.7**	16.7	14.0	**17.3**	**23.3**	20.7	**23.3**
NetEase-69	**7.3**	6.7	**7.3**	15.3	12.7	**16.0**	**21.3**	20.0	**21.3**
RTE	6.7	7.3	**8.0**	14.0	12.7	**14.7**	**22.0**	20.0	**22.0**

- **CA-greedy** selects the largest observed agreement rate Feng et al. (2022); our implementation uses its agreement signal as a scheduler.
- **Uncertainty Sampling** buys tuples for the edge with the largest current confidence radius, analogous in spirit to uncertainty sampling (Chen et al., 2025).
- **APT** uses the named Anytime Parameter-free Thresholding index of Locatelli et al. (2016). We retain it because it directly matches scalar threshold classification.
- **TaS-FG** is a UCB-style sampling rule adapted from the Track-and-Stop scheme (Russo et al., 2025). It prioritizes one-step reductions in the confidence radius relative to the remaining gap to the threshold.
- **ForcedEX** is a component variant that allocates 10% of its sampling budget to forced uncertainty-driven exploration to balance adaptivity with concentrated exploration (Maiti et al., 2026).

Metrics. The primary metric is certified-edge coverage; the secondary metric is false certification against the exact finite pool. Results are mean $\pm$ standard error over five seeds.

7.3 Main results and candidate-set stress

Table 1 shows the central result. BRCert has the highest mean coverage on all five datasets, exceeding margin-greedy by 0.6 percentage points on NetEase-126. Its largest improvements over margin-greedy are 4.0 points on CIFAR-10N and 2.0 points on Fact. TaS-FG is consistently weaker in this finite-budget regime, confirming that one-step uncertainty reduction does not target certificate completion.

Figure 2 identifies the intended regime without hiding dataset heterogeneity. With ten candidates, most promising pools can be substantially sampled and simple greedy baselines often tie BRCert. At 30 candidates BRCert wins or nearly ties across all datasets; at 50 candidates it produces particularly clear gains on CIFAR-10N and Fact. RTE at 50 edges is the main exception: many similar small binary pools make the estimate less discriminative.

Table 3: Finite-pool-cap sensitivity.

Dataset	$M = 15$			$M = 30$			$M = 45$		
	Margin	ForcedEX	BRCert	Margin	ForcedEX	BRCert	Margin	ForcedEX	BRCert
CIFAR-10N	52.0	51.3	**53.3**	24.0	26.0	**28.0**	18.0	**26.7**	**26.7**
Fact	44.0	39.3	**44.7**	20.0	18.7	**22.0**	12.7	12.7	**15.3**
NetEase-126	**33.3**	32.0	**33.3**	16.7	14.0	**17.3**	9.3	8.7	**10.0**
NetEase-69	21.3	**23.3**	21.3	15.3	12.7	**16.0**	**10.7**	10.0	**10.7**
RTE	32.0	31.3	**32.7**	14.0	12.7	**14.7**	**10.0**	**10.0**	**10.0**

Table 4: Component ablation. *(a) Dataset breakdown. (b) Candidate-set stress.*

(a) Variant	CIFAR-10N	Fact	NE-126	NE-69	RTE	Mean
TaS-FG	18.0	11.3	9.3	8.0	8.7	11.1
No reachability (Margin)	24.0	20.0	16.7	15.3	14.0	18.0
ForcedEX	26.0	18.7	14.0	12.7	12.7	16.8
Full BRCert	**28.0**	**22.0**	**17.3**	**16.0**	**14.7**	**19.6**

(b) Variant	10 edges	20 edges	30 edges	50 edges	Average
TaS-FG	48.4	20.4	11.1	6.2	21.5
No reachability (Margin)	**50.4**	27.8	18.0	10.9	26.8
ForcedEX	50.0	25.8	16.8	11.6	26.0
Full BRCert	**50.4**	**28.4**	**19.6**	**12.8**	**27.8**

7.4 Sensitivity and component ablations

Budget and pool size. Tables 2 and 3 report all completed sensitivity runs. At 30 edges and pool cap 30, increasing budget from 100 to 300 raises BRCert's mean coverage from 9.2% to 28.4% across datasets. With budget 200, reducing the cap from 30 to 15 raises mean coverage from 19.6% to 37.1% because smaller declared batches are easier to complete. Increasing the cap to 45 lowers mean coverage to 14.5%, with a decrease on every dataset, because the extra overlap does not offset the greater evidence needed to finish a certificate. Thus BRCert is most useful when the budget can finish some, but not all, candidate certificates.

Ablation. Table 4 reuses the candidate-set sweep to isolate BRCert's two design choices. Margin-greedy removes the reachability test; ForcedEX adds 10% forced uncertainty exploration; and TaS-FG replaces completion cost with one-step radius reduction. Full BRCert ties margin-greedy at 10 edges, then improves average coverage by 0.6, 1.6, and 1.9 percentage points at 20, 30, and 50 edges. Forced exploration is consistently harmful in this short-horizon setting, supporting the decision to use the shared history as the warm start and concentrate the purchase budget.

8 Conclusion

We introduced BRCert for federated data markets without verifiable ground truth. Rather than learning a new payment rule, budget-reachable certification asks which candidate certificate can still be completed with the tasks and budget that remain. This simple change matters under tight budgets: spreading evidence can leave every edge unresolved, whereas reachability-aware concentration converts overlap into usable certificates. Five real annotation markets show that BRCert is strongest when many heterogeneous edges compete for a budget that can finish only a subset. The contribution is deliberately narrow: a valid, lightweight assessment principle for fixed finite data batches.

AI USE STATEMENT

Generative AI tools assisted with language editing and code drafting. The authors verified and approved the research claims, proofs, citations, experiments, and final text.

ETHICS STATEMENT

This work uses existing annotation datasets and introduces no new human-subject data collection. Deployments should disclose the payment and deviation rules, protect retained reports, and provide an appeal or additional-evidence path when decisions affect suppliers.

REPRODUCIBILITY STATEMENT

The submission provides the manuscript source, aggregate experimental outputs, and the script used to regenerate its tables and figures. Appendix B records dataset handling and tuple construction, and Appendix A gives proof details.

REFERENCES

Anish Agarwal, Munther Dahleh, and Tuhin Sarkar. A marketplace for data: An algorithmic solution. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pp. 701–726, 2019.

Ayya Alieva, Ashok Cutkosky, and Abhimanyu Das. Robust pure exploration in linear bandits with limited budget. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139, pp. 187–195, 2021.

Abdellah Aznag, Robin Bowers, Rachel Cummings, Jason Hartline, Matthew vonAllmen, and Bo Waggoner. Sample complexity of peer prediction. *arXiv preprint arXiv:2608.16838*, 2026.

Rémi Bardenet and Odalric-Ambrym Maillard. Concentration inequalities for sampling without replacement. *Bernoulli*, 21(3):1361–1385, 2015.

Wenjing Chen, Shuo Xing, and Victoria G. Crawford. Adaptive threshold sampling for pure exploration in submodular bandits. In *Proceedings of the Forty-first Conference on Uncertainty in Artificial Intelligence*, volume 286, pp. 612–646, 2025.

Yiling Chen, Yiheng Shen, and Shuran Zheng. Truthful data acquisition via peer prediction. In *Advances in Neural Information Processing Systems*, volume 33, pp. 18194–18204, 2020.

Anirban Dasgupta and Arpita Ghosh. Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd International Conference on World Wide Web*, pp. 319–330, 2013.

Alexander P. Dawid and Allan M. Skene. Maximum likelihood estimation of observer error-rates using the em algorithm. *Journal of the Royal Statistical Society Series C*, 28(1):20–28, 1979.

Yiding Feng, Han Yu, and Yiling Chen. Peer prediction for learning agents. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Alice Gao, James R. Wright, and Kevin Leyton-Brown. Incentivizing evaluation via limited access to ground truth: Peer-prediction makes things worse. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pp. 567–584, 2016.

Amirata Ghorbani and James Zou. Data shapley: Equitable valuation of data for machine learning. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 2242–2251, 2019.

Steven R. Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. In *The Annals of Statistics*, volume 49, pp. 1055–1080, 2021.

Zehong Hu, Yitao Liang, Jie Zhang, Zhao Li, and Yang Liu. Inference aided reinforcement learning for incentive mechanism design in crowdsourcing. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Ruoxi Jia, David Dao, Boxin Wang, et al. Towards efficient data valuation based on the shapley value. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics*, pp. 1167–1176, 2019.

Radu Jurca and Boi Faltings. Mechanisms for making crowds truthful. *Journal of Artificial Intelligence Research*, 34:209–253, 2009.

Peter Kairouz, H. Brendan McMahan, Brendan Avent, et al. Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, 14(1–2):1–210, 2021.

Gabriella Kazai, Jaap Kamps, Marijn Koolen, and Natasa Milic-Frayling. Overview of the trec 2013 crowdsourcing track. In *Proceedings of the Twenty-Second Text REtrieval Conference*, 2013.

Yuqing Kong and Grant Schoenebeck. An information theoretic framework for designing information elicitation mechanisms that reward truth-telling. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pp. 129–148, 2019.

Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020.

Junnan Li, Chiyuan Zhang, Daniel Tarlow, and Richard Zemel. Crowdlab: Supervised learning to infer consensus labels and quality scores for data with multiple annotators. In *NeurIPS Data-Centric AI Workshop*, 2022.

Yang Liu and Yiling Chen. Sequential peer prediction: Learning to elicit effort using posted prices. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

Yang Liu and Yiling Chen. Surrogate scoring rules and a dominant truth serum for information elicitation. *arXiv preprint arXiv:1802.09158*, 2018.

Andrea Locatelli, Maurilio Gutzeit, and Alexandra Carpentier. An optimal algorithm for the thresholding bandit problem. In *Proceedings of the 33rd International Conference on Machine Learning*, pp. 1690–1698, 2016.

Arnab Maiti, Yunbei Xu, and Kevin Jamieson. On the power of adaptivity for ε-best arm identification in linear bandits. In *Proceedings of the 39th Conference on Learning Theory*, volume 336, pp. 4911–4968, 2026.

Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, pp. 1273–1282, 2017.

Nolan Miller, Paul Resnick, and Richard Zeckhauser. Eliciting informative feedback: The peer-prediction method. *Management Science*, 51(9):1359–1373, 2005.

Jian Pei. A survey on data pricing: From economics to data science. *IEEE Transactions on Knowledge and Data Engineering*, 34(10):4586–4608, 2022.

Drazen Prelec. A bayesian truth serum for subjective data. *Science*, 306(5695):462–466, 2004.

Goran Radanovic and Boi Faltings. Incentive schemes for participatory sensing. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1081–1089, 2015.

Vikas C. Raykar, Shipeng Yu, Linda H. Zhao, et al. Learning from crowds. *Journal of Machine Learning Research*, 11:1297–1322, 2010.

Alessio Russo, Yichen Song, and Aldo Pacchiano. Pure exploration with feedback graphs. In *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*, volume 258, pp. 1810–1818, 2025.

Robert J. Serfling. Probability inequalities for the sum in sampling without replacement. *The Annals of Statistics*, 2(1):39–48, 1974.

Victor Shnayder, Arpit Agarwal, Rafael Frongillo, and David C. Parkes. Informed truthfulness in multi-task peer prediction. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pp. 179–196, 2016.

Rachael Hwee Ling Sim, Yehong Zhang, Mun Choon Chan, and Bryan Kian Hsiang Low. Collaborative machine learning with incentive-aware model rewards. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 8927–8936, 2020.

Rion Snow, Brendan O'Connor, Daniel Jurafsky, and Andrew Y. Ng. Cheap and fast—but is it good? evaluating non-expert annotations for natural language tasks. In *Proceedings of EMNLP*, pp. 254–263, 2008.

Fei Wang, Haoyu Liu, Haoyang Bi, et al. A dataset for the validation of truth inference algorithms suitable for online deployment. *arXiv preprint arXiv:2403.08826*, 2024.

Juntao Wang, Yang Liu, and Yiling Chen. Forecast aggregation via peer prediction. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 9, pp. 131–142, 2021.

Ian Waudby-Smith and Aaditya Ramdas. Estimating means of bounded random variables by betting. *Journal of the Royal Statistical Society Series B*, 86(1):1–27, 2024.

Jiaheng Wei, Zhaowei Zhu, Hao Cheng, Tongliang Liu, Gang Niu, and Yang Liu. Learning with noisy labels revisited: A study using real-world human annotations. In *International Conference on Learning Representations*, 2022.

Jens Witkowski and David C. Parkes. Peer prediction without a common prior. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pp. 964–981, 2012.

Leon Witt, Togrul Abbasli, Kentaroh Toyoda, Wojciech Samek, and Lucy Klinger. Knowledge-free correlated agreement for incentivizing federated learning. *arXiv preprint arXiv:2605.04747*, 2026.

Hao Wu, Likun Zhang, Shucheng Li, Fengyuan Xu, and Sheng Zhong. Coast: Validation-free contribution assessment for federated learning based on cross-round valuation. *arXiv preprint arXiv:2409.02495*, 2024.

Yufeng Zhan, Peng Li, Zhihao Qu, Deze Zeng, and Song Guo. A survey of incentive mechanism design for federated learning. *IEEE Transactions on Emerging Topics in Computing*, 10(2):1035–1044, 2022.

Yudian Zheng, Guoliang Li, Yuanbing Li, Caihua Shan, and Reynold Cheng. Truth inference in crowdsourcing: Is the problem solved? In *Proceedings of the VLDB Endowment*, volume 10, pp. 541–552, 2017.

A Proofs

A.1 Proof of Theorem 1

Fix edge e, deviation f, and sample count n. Define $W_s = a_e[X_{e,s}(f^\star) - X_{e,s}(f)]$. Then $W_s \in [-2a_e, 2a_e]$ and is revealed without replacement from a population of size M_e. The Hoeffding–Serfling inequality bounds the deviation of the prefix mean from the finite-population mean by a range-four Hoeffding radius multiplied by $\sqrt{(M_e - n)/(M_e - 1)}$. Assigning failure probability $\delta/(|E|(|\mathcal{F}| - 1)M_{\max})$ to every (e, f, n) gives Eq. (5). A union bound yields a single event of probability at least $1 - \delta$ covering all declared edges, deviations, and sample counts. Selecting an edge adaptively only selects one member of this already simultaneous family. Taking a minimum over deviations gives the claimed bound for Γ_e.

A.2 Proof sketch of Proposition 1

By definition, $m_e(t)$ is the first future count for which the plug-in lower bound crosses ε. If the empirical deviation margins remain fixed and BRCert supplies these observations, the radius follows exactly the sequence used by the reachability calculation, so certification occurs no later than that count. Theorem 1 concerns the actual observed prefix and therefore remains valid even when the stability premise fails; only the scheduling prediction, not certificate validity, is affected.

A.3 Closed-form reachability approximation

Ignoring integer rounding and writing $C = 8\log(2|E|(|\mathcal{F}|-1)M_{\max}/\delta)$, the radius at future count n is

$$r_e(n)^2 = a_e^2 C \frac{M_e - n}{n(M_e - 1)}. \tag{9}$$

For a positive plug-in slack $d_e = \widehat{\Gamma}_e - \varepsilon > 0$, the inequality $r_e(n) \le d_e$ is equivalent to

$$n \ge n_e^\star := \frac{a_e^2 C M_e}{d_e^2(M_e - 1) + a_e^2 C}. \tag{10}$$

Thus the continuous approximation to the additional evidence requirement is

$$\widetilde{m}_e(t) = \max\{0, \lceil n_e^\star \rceil - n_e\}. \tag{11}$$

The implementation evaluates the discrete radius exactly, but Eq. (10) makes the mechanism transparent. Required evidence decreases with the squared empirical slack d_e^2, increases with payment scale a_e^2, and is capped by the finite pool. In the large-pool limit, $n_e^\star \to a_e^2 C/d_e^2$, recovering the familiar inverse-square dependence. When n approaches M_e, the finite-population term vanishes and the declared pool becomes exactly known.

A.4 Reachability, infeasibility, and budget accounting

Define the step-t reachable set

$$\mathcal{R}_t = \{e : \widehat{\Gamma}_e > \varepsilon,\ m_e(t) \le R_e,\ c_e m_e(t) \le B_t\}. \tag{12}$$

The first condition rules out a nonpositive plug-in slack; the second prevents requesting more tuples than remain on the edge; and the third imposes the global purchasing horizon. The selected edge is

$$e_t \in \arg\min_{e \in \mathcal{R}_t} c_e m_e(t), \tag{13}$$

with empirical margin used only for tie breaking. This is a myopic certificate-completion rule, not a claim of globally optimal knapsack allocation. Nevertheless, if the stability premise of Proposition 1 holds for the chosen edge, the realized completion cost is at most the predicted $c_e m_e(t)$, so the step cannot violate the posted budget.

For intuition, suppose all costs are one and a set S of reachable certificates has stable requirements m_e. Any policy that completes every edge in S needs at least $\sum_{e \in S} m_e$ observations. Sorting by m_e therefore maximizes the number completed under a fixed budget in this deterministic surrogate. BRCert repeatedly solves this surrogate with updated empirical margins. The argument explains its concentration behavior while also clarifying the limitation: margin changes can reorder the surrogate costs.

A.5 Monotonicity of the finite-pool radius

For $1 \le n < M_e$, Eq. (9) gives

$$\frac{r_e(n+1)^2}{r_e(n)^2} = \frac{n(M_e - n - 1)}{(n+1)(M_e - n)} < 1. \tag{14}$$

Hence the radius is strictly decreasing until pool exhaustion, which makes the first-crossing definition of $m_e(t)$ well posed. It also exposes the difference from a one-step UCB heuristic. The latter ranks the immediate decrement $r_e(n) - r_e(n+1)$, whereas BRCert ranks the full remaining distance to a valid threshold crossing. Large immediate uncertainty reduction need not imply affordable certificate completion.

A.6 Validity under adaptive edge choice

Let $\mathcal{E}_0$ be the simultaneous event on which every confidence statement for every edge, deviation, and prefix holds. The proof of Theorem 1 gives $\Pr(\mathcal{E}_0) \geq 1 - \delta$ before any scheduling decisions. For any adaptive history H_t, the chosen edge $e_t = g_t(H_t)$ merely indexes one inequality already contained in $\mathcal{E}_0$. Therefore

$$\Pr(\exists t : \mathrm{LCB}_{e_t}(n_{e_t}) > \Gamma_{e_t}) \leq \Pr(\mathcal{E}_0^c) \leq \delta. \tag{15}$$

The reachability rule may affect which valid statements are observed and how many certificates are completed, but it does not change their simultaneous error control. This separation is why the same certificate can be compared across all schedulers.

B Detailed Experimental Protocol

B.1 Dataset handling and tuple construction

We remove duplicate reports for the same worker and task, map labels to consecutive integers, and retain worker pairs with enough common items to form the requested pool. Each KFCA tuple uses one shared bonus item and two distinct penalty items. Candidate selection uses overlap only; report values are not used to choose the candidate graph. Within each edge, a seeded random permutation fixes the history and purchase order before any scheduler runs. Thus every scheduler sees identical finite pools for a given seed.

Fact and RTE use their public crowd labels. The two NetEase subsets are identified by market identifiers 69 and 126. For CIFAR-10N, the three human annotations associated with each image are treated as reports; the image pixels and trained classifiers are never used. Gold labels, when available, are used only for the descriptive worker-accuracy statistic and never enter payment, scheduling, or certification.

All reported experiments use seeds 0–4. The candidate-edge sweep uses 10, 20, 30, and 50 edges at budget 200 and cap 30. The budget sweep uses 100, 200, and 300 at 30 edges and cap 30. The pool sweep uses caps 15, 30, and 45 at 30 edges and budget 200. Across this complete matrix, the exact finite-pool audit records no false certificate.

C Additional Discussion

Why budget reachability rather than UCB or reinforcement learning? The decision variable is evidence allocation under a short, known horizon. A one-step UCB score values uncertainty reduction but need not finish a certificate; reinforcement learning adds training and approximation without changing certificate validity. BRCert directly targets the operational bottleneck, namely whether an edge can cross the valid threshold before its pool or budget ends.

Why not choose the most similar peer? Raw agreement conflates informative correlation with shared bias or coordinated constant reporting. Margin-greedy is stronger, but it can keep sampling a promising edge that is still impossible to certify within the remaining horizon. BRCert adds only this reachability test.

D Limitations and Broader Impact

The certificate concerns a declared deviation class and finite batch under the assumptions of the underlying peer-payment rule; it is not a semantic quality label. Correlated low-quality reports and coalitions remain possible, and concentrated allocation may leave atypical suppliers unresolved.