Calibrating Mechanisms with Delayed Evidence

Yuchen Hu
Columbia University

Stephen Bates
MIT

Martin J. Wainwright
MIT

1 Introduction and Problem Setup

Many operational decisions need to be made before reliably assessing all relevant information. For example, an AI developer may need to choose a data provider before knowing how informative the resulting data will be for training. While a provider's cost always remains private, the quality or operational impact of what it provides can usually be assessed from data generated after the decision is made. If known upfront, standard procurement mechanisms could be used, but asking the provider to report it creates a new strategic incentive that such mechanisms were not designed to address.

We study a simple alternative to redesigning the mechanism from scratch. Using agents' reports, the principal runs the mechanism that would be used if the operational impact were known, and upon observing noisy post-decision evidence, corrects their payoffs. We provide a calibrated correction that protects truthful agents against erroneous correction while disciplining reports that are inconsistent with later evidence. The resulting incentive loss is controlled directly by the calibration error and statistical resolution of the verifier, and vanishes as the evidence becomes more informative.

This complements work on data-driven mechanism design and on ex-post or probabilistic verification [Mylovanov and Zapechelnyuk, 2017, Alaei et al., 2024, Bergemann et al., 2024, Ball and Kattwinkel, 2025, Ben-Porath et al., 2026]. Instead of verifying the agent's private cost or value or redesigning mechanism-specific transfers, we use delayed evidence about auxiliary operational characteristics as a modular correction layer. It leaves the underlying mechanism unchanged and can be used with a broad range of verification procedures including concentration-based tests and anytime-valid e-processes. We illustrate the framework for standard resource-allocation mechanisms in Section A.

Oracle benchmark. Consider n strategic agents. Agent i privately observes $\theta_i = (\theta_i^{\mathrm{p}}, \theta_i^{\mathrm{e}}) \in \Theta_i^{\mathrm{p}} \times \Theta_i^{\mathrm{e}}$. The private type θ_i^{p} may remain unobserved throughout the interaction, while the environmental type θ_i^{e} is an additional payoff-relevant characteristic that can be learned from post-action data. For example, in data procurement, θ_i^{p} can be a provider's production cost and θ_i^{e} its data informativeness. If informativeness were known, the principal could run a standard cost-DSIC procurement mechanism, such as a Myerson-type allocation rule [Myerson, 1981].

We take as given such an oracle benchmark mechanism $m^*(\hat{\theta}^{\mathrm{p}}, \theta^{\mathrm{e}})$ that would be used if the environmental profile were observed before the principal acts. For every fixed θ^{e}, the benchmark is DSIC and IR in the private type: for every agent i, opponents' reports $\hat{\theta}_{-i}^{\mathrm{p}}$, and alternative report $\hat{\theta}_i^{\mathrm{p}}$,

$$U_i^*((\theta_i^{\mathrm{p}}, \theta_i^{\mathrm{e}}), \widehat{\theta}_{-i}; \theta_i) \geq U_i^*((\hat{\theta}_i^{\mathrm{p}}, \theta_i^{\mathrm{e}}), \widehat{\theta}_{-i}; \theta_i), \tag{1}$$

where U_i^* denotes agent i's utility under the mechanism m^*, and truthful reporting gives nonnegative utility. Such benchmarks are standard in many single-parameter settings; we give VCG and Myerson examples in Section A in a resource-allocation setting.

Delayed verification and correction. At the start of the game, each agent is asked to report both components, $\widehat{\theta}_i = (\hat{\theta}_i^{\mathrm{p}}, \hat{\theta}_i^{\mathrm{e}})$, and the principal takes the benchmark action evaluated at the reported environmental profile, $m^*(\hat{\theta}^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}})$. Only after this action is taken does the interaction generate evidence X informative about θ^{e}. The principal may then apply an ex-post correction, such as withholding payment, terminating a contract, or imposing an additional service requirement.

40th Conference on Neural Information Processing Systems (NeurIPS 2026). Workshop: Second Workshop on ML×OR.

For a report profile $\widehat{\theta}$, recall that $U_i^*(\widehat{\theta};\theta_i)$ denotes agent i's utility when the benchmark outcome is left uncorrected, and let $U_i^-(\widehat{\theta};\theta_i)$ denote its utility under the strongest feasible ex-post correction. Define the strongest correction gap $\Delta_i(\widehat{\theta};\theta_i) := U_i^*(\widehat{\theta};\theta_i) - U_i^-(\widehat{\theta};\theta_i)$. We impose the following assumption on the strength of ex-post correction.

Assumption 1 (Sufficient correction technology)**.** There exist constants $\kappa > 0$ and $\overline{\Delta} < \infty$ such that

$$U_i^*(\theta_i, \widehat{\theta}_{-i};\theta_i) - U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i};\theta_i) \geq \kappa \tag{2}$$

and $0 \leq \Delta_i(\widehat{\theta};\theta_i) \leq \overline{\Delta}$.

This condition requires the available ex-post action to be strong enough to eliminate the gain from a false report. With exact ex-post verification, Assumption 1 ensures that maximal correction makes every false environmental report strictly worse than the truthful oracle outcome; leaving truthful environmental reports uncorrected then preserves the benchmark incentives in the private type.

2 Main Results

2.1 Calibrated correction for noisy verification

We now consider the case in which the environmental type can only be learned imperfectly from post-action evidence. To formalize increasingly informative verification, we study a sequence of problem instances $k = 1, 2, ...$ The index k may represent a growing sample size, a longer observation horizon, or more generally a scaling of the underlying decision problem that generates increasingly informative evidence. This allows us to characterize the finite-sample and asymptotic incentives of agents in terms of the general statistical resolution of the delayed evidence. Throughout, we suppress the dependence of quantities on k when it is not essential.

After the benchmark action $m^*(\widehat{\theta})$ is taken in instance k, evidence X_k is generated about the agents' environmental types. For each agent i, let $\phi_{ik}(\widehat{\theta}, X_k) \in [0,1]$ denote the extent to which the reported environmental type passes verification. In the binary case, $\phi_{ik} = 1$ corresponds to passing an ex-post verification and $\phi_{ik} = 0$ to maximal correction due to failure.

We summarize the quality of the verification procedure below through two quantities. The calibration error α_k controls erroneous correction of a truthful environmental report, while s_k measures the statistical resolution at which deviations in the environmental type become detectable. The test is calibrated so that both errors become negligible as the evidence becomes more informative.

Assumption 2 (Calibrated verification)**.** There exist sequences $\alpha_k, s_k \to 0$, and a fixed nonincreasing function $\rho : [0,\infty) \to [0,1]$, $\rho(t) \to 0$ as $t \to \infty$, such that, whenever post-action evidence X_k is generated,[1]

$$\mathbb{E}_{\theta_i^{\mathrm{e}}}\left[\phi_{ik}(\widehat{\theta}, X_k)\right] \geq 1 - \alpha_k, \qquad \text{if } \hat{\theta}_i^{\mathrm{e}} = \theta_i^{\mathrm{e}}, \tag{3a}$$

and

$$\mathbb{E}_{\theta_i^{\mathrm{e}}}\left[\phi_{ik}(\widehat{\theta}, X_k)\right] \leq \rho\left(\left\|\hat{\theta}_i^{\mathrm{e}} - \theta_i^{\mathrm{e}}\right\| / s_k\right). \tag{3b}$$

Assumption 2 requires that the verification is calibrated so that a truthful environmental report has expected verification score at least $1 - \alpha_k$, under which a deviation that is large relative to s_k has increasingly small expected score. The sequence s_k therefore captures the effective statistical resolution of the verifier at the designed calibration level α_k. In Section 2.3, we give examples of s_k for concentration-based verification and for sequential verification based on e-processes, and discuss the choice of the design parameter α_k as well as the economic cost of correction.

Given such a verifier, we define the calibrated correction mechanism by giving the agent a payoff of

$$\hat{U}_i(\widehat{\theta};\theta_i, X_k) := U_i^-(\widehat{\theta};\theta_i) + \frac{\phi_{ik}(\widehat{\theta}, X_k)}{1-\alpha_k}\left(U_i^*(\widehat{\theta};\theta_i) - U_i^-(\widehat{\theta};\theta_i)\right). \tag{4}$$

When verification fails completely, the agent receives the maximally corrected payoff U_i^-. When verification succeeds, the adjustable portion of the payoff is scaled by $1/(1-\alpha_k)$. This adjustment compensates for the statistical risk of erroneous correction under truthful reporting.

[1] We assume that any report yielding positive benchmark utility generates evidence X_k; see Section B for the convention.

2.2 Approximate incentive recovery

We now characterize when calibrated correction recovers the incentive properties of the oracle benchmark. For an environmental report $\hat{\theta}_i^{\mathrm{e}}$ of agent i, define the agent's benchmark utility at $\hat{\theta}_i^{\mathrm{e}}$ as

$$B_i(\hat{\theta}_i^{\mathrm{e}}) := \sup_{\hat{\theta}_i^{\mathrm{p}}} U_i^*((\hat{\theta}_i^{\mathrm{p}}, \hat{\theta}_i^{\mathrm{e}}), \widehat{\theta}_{-i}; \theta_i), \tag{5}$$

where we suppress the dependence of B_i on $(\widehat{\theta}_{-i}, \theta_i)$. This is the optimal utility the agent can achieve under the benchmark mechanism fixing one (potentially nontruthful) environmental report. We make the following one-sided local Lipschitz condition on this benchmark utility.

Assumption 3 (Benchmark stability)**.** There exist constants $L_B < \infty$ and $r_0 > 0$ such that, uniformly over agents, types, and opponents' reports, for every environmental report $\hat{\theta}_i^{\mathrm{e}}$ satisfying $\left\|\hat{\theta}_i^{\mathrm{e}} - \theta_i^{\mathrm{e}}\right\| \le r_0$,

$$B_i(\hat{\theta}_i^{\mathrm{e}}) - B_i(\theta_i^{\mathrm{e}}) \le L_B \left\|\hat{\theta}_i^{\mathrm{e}} - \theta_i^{\mathrm{e}}\right\|.^2 \tag{6}$$

Assumption 3 says that a small perturbation of the environmental input cannot generate a disproportionate increase in the agent's best attainable benchmark utility. Below, we show that under this assumption, calibrated correction asymptotically recovers the incentive guarantees of the oracle mechanism.

Theorem 1. *Under Assumption 1–Assumption 3, for all sufficiently large k, agent i, type θ_i, and realized profile of opponents' reports $\widehat{\theta}_{-i}$,*

$$\sup_{\widehat{\theta}_i} \mathbb{E}\left[\hat{U}_i(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] - \mathbb{E}\left[\hat{U}_i(\theta_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] = \mathcal{O}\left(s_k + \alpha_k\right). \tag{7}$$

Moreover, truthful reporting gives nonnegative expected utility.

Theorem 1 gives a pointwise incentive guarantee where truthful reporting of both the private and environmental types is an ϵ_k-best response for every realized profile of opponents' reports. Thus, the calibrated mechanism is asymptotically dominant-strategy incentive compatible, and under any prior over agent types, truthful reporting also forms an ϵ_k-Bayesian Nash equilibrium. In Section B, we further characterize exact best responses under a margin condition on the benchmark utility. With a β-margin, every active exact best response lies within an $\mathcal{O}(\epsilon_k^{1/\beta})$ neighborhood of the truthful type.

Remark 1. In Section A.1, we show that Assumption 3 can hold globally for broad classes of resource-allocation mechanisms even when the optimal allocation itself jumps at decision boundaries. What we require is only that a small perturbation of the environmental type cannot change the objective used to choose the allocation by too much. In mechanisms where crossing a decision boundary can instead cause a discrete utility jump, an analogous condition in expectation over other agents' types may still yield Bayesian incentive guarantees, which we leave for future work.

2.3 Statistical resolution and the cost of calibrated correction

Above, we showed that the incentive error is $\mathcal{O}(s_k + \alpha_k)$, where the calibration level α_k chosen by the principal and the resulting verification resolution s_k generally trade off: reducing α_k protects truthful reports from erroneous correction, but requires a more lenient verifier and hence a coarser resolution. It is therefore natural to calibrate the verifier so that the two errors are balanced.

Concentration-based verification. Suppose the post-action evidence X_k yields an estimator $\tilde{\theta}_{ik}^{\mathrm{e}}$ satisfying the sub-Gaussian bound $\mathcal{P}_{\theta_i^{\mathrm{e}}}[||\tilde{\theta}_{ik}^{\mathrm{e}} - \theta_i^{\mathrm{e}}|| > r_k\, t] \le c_1 \exp(-c_2 t^2)$, $t > 0$, for some estimation rate $r_k \to 0$. Consider a simple binary verifier $\phi_{ik} = I\left\{||\tilde{\theta}_{ik}^{\mathrm{e}} - \hat{\theta}_{ik}^{\mathrm{e}}|| \le s_k\right\}$. Under the sub-Gaussian bound, a truthful report is rejected with probability at most $\alpha_k := c_1 \exp(-c_2 s_k^2 / r_k^2)$. Choosing the calibration so that $\alpha_k \asymp s_k$ gives

$$\alpha_k, s_k = \mathcal{O}\left(r_k \sqrt{\log(1/r_k)}\right),$$

[2]The linear Lipschitz bound can be replaced by any modulus of continuity $\omega(||\hat{\theta}_i^{\mathrm{e}} - \theta_i^{\mathrm{e}}||) \to 0$, which, in turn, will yield a corresponding incentive bound $\mathcal{O}(\omega(s_k) + \alpha_k)$.

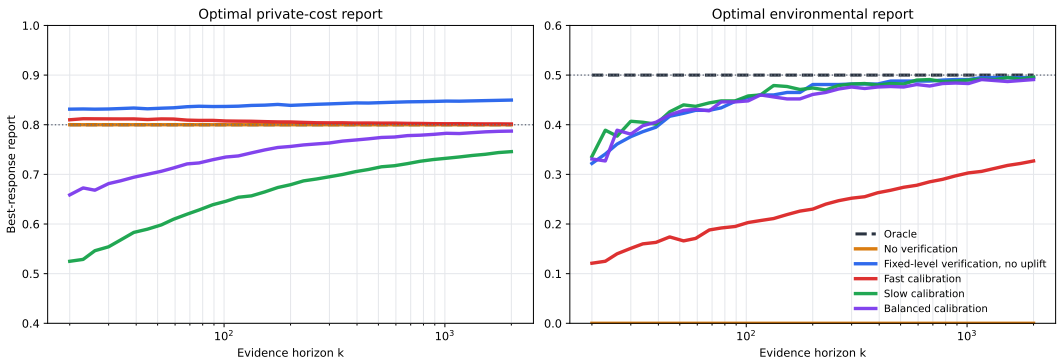


Figure 1: Best-response reports under staged e-process verification.

and therefore $\epsilon_k = \mathcal{O}(r_k\sqrt{\log(1/r_k)})$. With the standard parametric rate $r_k = \mathcal{O}(k^{-1/2})$, this becomes $\mathcal{O}(\sqrt{\log k/k})$.

Sequential verification with e-processes. Suppose evidence arrives sequentially over stages $l = 1, ..., k$, and let $E_{il}(\hat{\theta}_i^{\mathrm{e}})$ be an e-process testing the reported environmental type [Ramdas et al., 2023]. Such e-processes yield the anytime-valid bound $\mathbb{P}_{\theta_i^{\mathrm{e}}}\left[\sup_{l\le k} E_{il}(\theta_i^{\mathrm{e}}) \ge 1/\alpha_k\right] \le \alpha_k$.

Consider a staged contract that terminates when $E_{il}(\hat{\theta}_i^{\mathrm{e}})$ first exceeds $1/\alpha_k$. Upon termination, all future payments are withheld; before termination, each stage of the adjustable benchmark payment is uplifted by $1/(1-\alpha_k)$. Thus, if ϕ_{ik} denotes the fraction of the staged payoff retained before termination, the realized utility is exactly the calibrated correction in Equation (4).

Under a standard local quadratic log-growth condition, suppose a misspecified report at distance δ accumulates log-evidence at rate $R_k\delta^2/k$ per stage, where $R_k \to \infty$ denotes the effective information over k stages. Its threshold-crossing time is then of order $k\log(1/\alpha_k)/(R_k\delta^2)$, resulting in a verification resolution $s_k = \mathcal{O}(\sqrt{\log(1/\alpha_k)/R_k})$. Balancing $\alpha_k \asymp s_k$ yields

$$\alpha_k, s_k, \epsilon_k = \mathcal{O}\left(\sqrt{\log R_k/R_k}\right).$$

The calibration level also determines the economic cost of the correction. Under truthful environmental reporting and suppressing the common arguments $((\hat{\theta}_i^{\mathrm{p}}, \theta_i^{\mathrm{e}}), \widehat{\theta}_{-i}; \theta_i)$,

$$\mathbb{E}\left[\hat{U}_i\right] - U_i^* = \frac{\mathbb{E}\left[\phi_{ik}\right] - (1-\alpha_k)}{1-\alpha_k}\Delta_i \in \left[0, \frac{\alpha_k}{1-\alpha_k}\Delta_i\right]. \tag{8}$$

Thus, exact calibration $\mathbb{E}\left[\phi_{ik}\right] = 1-\alpha_k$ preserves benchmark utility exactly, while with bounded Δ_i the expected uplift vanishes as $\alpha_k \to 0$. In payment-based implementations, this uplift corresponds directly to the principal's additional expected payment.

3 Numerical Illustration

We illustrate the incentive effects of calibrated correction using a service-procurement example detailed in Section C. Post-allocation evidence arrives sequentially, and verification uses the Gaussian mixture e-process with staged payments [Howard et al., 2021]. Figure 1 illustrates the resulting tradeoff. Without verification, the private cost is reported truthfully but the environmental type is maximally understated. Fixed-level verification increasingly disciplines the environmental report, but the absence of an uplift leaves a persistent distortion in the private-cost report. Among calibrated mechanisms, a rapidly vanishing α_k largely preserves the private-cost incentive but raises the verification threshold too rapidly, making verification overly lenient and slowing recovery of the environmental type. Conversely, a slowly vanishing α_k yields sharper environmental verification but a larger finite-sample uplift, generating more distortion in the private-cost report. The balanced choice avoids both extreme and brings both reports progressively toward truth.

References

Saeed Alaei, Alexandre Belloni, Ali Makhdoumi, and Azarakhsh Malekian. Optimal auction design with deferred inspection and reward. *Operations Research*, 72(6):2413–2429, 2024.

Ian Ball and Deniz Kattwinkel. Probabilistic verification in mechanism design. *Theoretical Economics*, 20(4):1247–1284, 2025.

Elchanan Ben-Porath, Eddie Dekel, and Barton L Lipman. Mechanism design for acquisition of/stochastic evidence. *Theoretical Economics*, 21(1):99–131, 2026.

Dirk Bergemann, Marek Bojko, Paul Dütting, Renato Paes Leme, Haifeng Xu, and Song Zuo. Data-driven mechanism design: Jointly eliciting preferences and information. *arXiv preprint arXiv:2412.16132*, 2024.

Edward H. Clarke. Multipart pricing of public goods. *Public choice*, 11(1):17–33, 1971.

Theodore Groves. Incentives in teams. *Econometrica*, 41(4):617–631, 1973.

Steven R Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2):1055–1080, 2021.

Roger B. Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

Tymofiy Mylovanov and Andriy Zapechelnyuk. Optimal allocation with ex post verification and limited penalties. *American Economic Review*, 107(9):2666–2694, 2017.

Aaditya Ramdas, Peter Grünwald, Vladimir Vovk, and Glenn Shafer. Game-theoretic statistics and safe anytime-valid inference. *Statistical Science*, 38(4):576–601, 2023.

William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37, 1961.

A Resource Allocation with Verifiable Resource Intensity

Consider a resource allocation problem with feasible allocations $q = (q_1, ..., q_n)$ in a compact set $\mathcal{Q} \subset [0,1]^n$ containing the zero allocation. Agent i's private type $\theta_i^{\mathrm{p}} \in [\underline{\theta^{\mathrm{p}}}, \overline{\theta^{\mathrm{p}}}]$ is its value for receiving the resource, so that its value from allocation q is $\theta_i^{\mathrm{p}} q_i$. The environmental profile θ^{e} determines the principal-side value of an allocation through a function $F(q, \theta^{\mathrm{e}})$. For example, θ_i^{e} may describe the resource intensity or congestion generated by serving agent i. The data-procurement example is the analogous procurement setting, where q_i indicates purchasing from provider i, θ_i^{p} is its private cost, and $F(q, \theta^{\mathrm{e}})$ captures the value of the acquired data. The exact formulation can be obtained by reversing the sign of the agent-side term.

We consider the following two benchmark mechanisms that target slightly different objectives.

Efficient benchmark. Suppose the goal is to maximize the total social welfare,

$$q^{VCG}(\hat{\theta}^{\mathrm{p}}, \theta^{\mathrm{e}}) \in \operatorname{argmax}_{q \in \mathcal{Q}} \left\{ F(q, \theta^{\mathrm{e}}) + \sum_i \hat{\theta}_i^{\mathrm{p}} q_i \right\}. \tag{9a}$$

Together with the standard VCG transfers [Vickrey, 1961, Clarke, 1971, Groves, 1973], this allocation is DSIC and IR in the agents' private values. However, this guarantee treats the environmental profile as known. If agent i were also allowed to report $\hat{\theta}_i^{\mathrm{e}}$, its report would change the system-side objective $F(q, \theta^{\mathrm{e}})$ rather than its own valuation $\hat{\theta}_i^{\mathrm{p}} q_i$. Standard VCG payments therefore do not in general make truthful reporting of θ_i^{e} incentive compatible.

Myerson benchmark. Suppose instead that the principal maximizes its expected payoff including both the principal-side value and the collected payments. Let the private value θ_i^{p} be independently distributed according to CDF G_i with density g_i, and define its virtual value

$$\psi_i(v) = v - \frac{1 - G_i(v)}{g_i(v)}.$$

Under the usual regularity condition that ψ_i is increasing, the Myerson benchmark allocates according to

$$q^M(\hat{\theta}^{\mathrm{p}}, \theta^{\mathrm{e}}) \in \operatorname{argmax}_{q \in \mathcal{Q}} \left\{ F(q, \theta^{\mathrm{e}}) + \sum_i \psi_i(\hat{\theta}_i^{\mathrm{p}}) q_i \right\}. \tag{9b}$$

Together with the standard single-parameter Myerson payments [Myerson, 1981], this benchmark is DSIC and IR in the private values. As with the VCG benchmark, this incentive guarantee holds for a fixed environmental profile. When θ_i^{e} is also privately reported and enters F, the standard single-parameter payment characterization does not by itself provide incentives for truthful reporting of θ_i^{e}.

A.1 Stability of the allocation benchmarks

For both benchmarks, the optimal allocation q may change discontinuously as the environmental profile varies, for example near allocation ties. Nevertheless, our next results show that Assumption 3 continues to hold under a simple stability condition on the primitive system-side objective $F(q, \theta^{\mathrm{e}})$.

We impose a regularity condition directly on the system-side objective. Suppose there exists a constant $L_F < \infty$ such that, for every feasible allocation q, agent i, opponents' environmental profile $\hat{\theta}^{\mathrm{e}}_{-i}$, and $e, e' \in \Theta^{\mathrm{e}}$,

$$\left| F\left(q, (e, \hat{\theta}^{\mathrm{e}}_{-i})\right) - F\left(q, (e', \hat{\theta}^{\mathrm{e}}_{-i})\right) \right| \leq L_F \left\| e - e' \right\|. \tag{10}$$

Thus, for any fixed allocation, a small change in an agent's environmental characteristic produces only a proportional change in system value, while the optimal allocation itself need not vary continuously with e.

For either benchmark, define agent's benchmark utility B_i as in Equation (5). We show below that Assumption 3 holds for both benchmarks.

Proposition 2. *Suppose F satisfies Equation* (10)*. Then, the VCG benchmark satisfies*

$$\left| B_i^{VCG}(e) - B_i^{VCG}(e') \right| \leq 2L_F \left\| e - e' \right\|$$

uniformly over agents, types, and opponents' reports.

Proposition 3. *Suppose F satisfies Equation* (10)*, and each virtual-value function ψ_i is twice continuously differentiable with strictly positive derivative. Then, there exists a constant $L_M < \infty$ such that*

$$\left| B_i^{M}(e) - B_i^{M}(e') \right| \leq L_M \left\| e - e' \right\|$$

uniformly over agents, types, and opponents' reports.

B Exact Best Responses under a Margin Condition

Theorem 1 shows that truthful reporting is ϵ_k-optimal under calibrated correction, and no deviation can improve the agent's expected utility by more than $\epsilon_k = \mathcal{O}(s_k + \alpha_k)$. This result does not, however, characterize the reports that exactly maximize utility under the calibrated mechanism. In this section, we show under an additional margin condition on the benchmark mechanism that every exact best response must itself converge to the truthful type.

Assumption 4 (Private-type margin)**.** There exist constants $C_M > 0$, $r_0 > 0$,[3] and $\beta \geq 1$ such that, uniformly over agents, types, and opponents' reports, for every environmental report $\hat{\theta}_i^{\mathrm{e}}$ satisfying $\left\| \hat{\theta}_i^{\mathrm{e}} - \theta_i^{\mathrm{e}} \right\| \leq r_0$,

$$U_i^*((\theta_i^{\mathrm{p}}, \hat{\theta}_i^{\mathrm{e}}), \widehat{\theta}_{-i}; \theta_i) - U_i^*((\hat{\theta}_i^{\mathrm{p}}, \hat{\theta}_i^{\mathrm{e}}), \widehat{\theta}_{-i}; \theta_i) \geq C_M \left(\left\| \hat{\theta}_i^{\mathrm{p}} - \theta_i^{\mathrm{p}} \right\| \wedge r_0 \right)^{\beta}. \tag{11}$$

[3] Without loss of generality, we take r_0 to be the smaller of the neighborhood radii required by Assumption 3 and Assumption 4.

Assumption 4 quantifies how strongly the benchmark mechanism identifies the truthful private report. In smooth interior problems, one typically obtains $\beta = 2$ from local curvature of the benchmark utility, while smaller values of β correspond to a sharper loss from private-type misreport.

We next characterize the exact optimal reports under calibrated correction. In some benchmark mechanisms, only selected agents enter the post-action interaction and generate verification evidence. Given opponents' reports $\widehat{\theta}_{-i}$, we call a report $\widehat{\theta}_i$ active if the benchmark action under $(\widehat{\theta}_i, \widehat{\theta}_{-i})$ induces such an interaction for agent i. Since calibrated correction leaves the benchmark pre-data action unchanged, the same active reports generate evidence under the calibrated mechanism.

Proposition 4. *Under Assumption 1–Assumption 4, for all sufficiently large k, agent i, type θ_i, realized profile of opponents' reports $\widehat{\theta}_{-i}$, and every active exact best response $\widehat{\theta}_{ik} = (\hat{\theta}^p_{ik}, \hat{\theta}^e_{ik})$ of agent i,*

$$\left\|\hat{\theta}^e_{ik} - \theta^e_i\right\| = \mathcal{O}\left(s_k\right), \qquad \left\|\hat{\theta}^p_{ik} - \theta^p_i\right\| = \mathcal{O}\left((s_k + \alpha_k)^{1/\beta}\right). \tag{12}$$

Proposition 4 complements the approximate incentive guarantee in Theorem 1. Not only is truthful reporting approximately optimal, but every active exact best response is itself approximately truthful. In particular, the environmental report is within the statistical resolution s_k of the truth, while the private-report error is $\mathcal{O}\left((s_k + \alpha_k)^{1/\beta}\right)$, with β capturing the curvature of the benchmark incentive problem.

Assumption 4 and Proposition 4 above assume for simplicity that truthful reporting is the unique optimal private-type report in the benchmark mechanism. When the benchmark is only weakly DSIC, there may be other private reports that yield the same benchmark utility. In this case, Assumption 4 and Proposition 4 can instead be stated in terms of the distance to the benchmark best-response set. The same argument then shows that every exact best response under calibrated correction lies within $\mathcal{O}\left((s_k + \alpha_k)^{1/\beta}\right)$ of this set.

For discrete private types, the result can be strengthened to exact recovery of the private type. If distinct private types are separated by a fixed positive distance, then $s_k + \alpha_k \to 0$ implies that, for all sufficiently large k, every exact best response satisfies $\hat{\theta}^{\mathrm{p}}_{ik} = \theta^{\mathrm{p}}_i$.

C Details of Numerical Illustration

We illustrate the incentive effects of calibrated correction in a simple single-agent resource-allocation problem. The agent has type $\theta = (\theta^{\mathrm{p}}, \theta^{\mathrm{e}}) \in [0,1]^2$, where θ^{p} is the agent's private marginal cost of providing the service, and θ^{e} is an environmental/operational cost imposed on the principal or system when using this provider, and it can be learned from execution data. Given reports $(\hat{\theta}^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}})$, consider the oracle allocation rule

$$q^*(\hat{\theta}^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}}) = 1 - \frac{\hat{\theta}^{\mathrm{p}} + \hat{\theta}^{\mathrm{e}}}{2}.$$

For any fixed $\hat{\theta}^{\mathrm{e}}$, this rule is monotone in the reported cost. Normalizing the highest-cost type's utility to zero, the standard single-parameter payment identity gives

$$\begin{aligned} P^*(\hat{\theta}^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}}) &= \hat{\theta}^{\mathrm{p}} q^*(\hat{\theta}^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}}) + \int_{\hat{\theta}^{\mathrm{p}}}^{1} q^*(z, \hat{\theta}^{\mathrm{e}}) dz \\ &= \frac{3}{4} - \frac{\hat{\theta}^{\mathrm{e}}}{2} - \frac{(\hat{\theta}^{\mathrm{p}})^2}{4}. \end{aligned}$$

Thus, absent correction, the agent's utility is $U^*(\hat{\theta}^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}}; \theta) = P^*(\hat{\theta}^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}}) - \theta^{\mathrm{p}} q^*(\hat{\theta}^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}})$. For every fixed environmental report $\hat{\theta}^{\mathrm{e}}$,

$$U^*(\theta^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}}; \theta) - U^*(\hat{\theta}^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}}; \theta) = \frac{1}{4}\left(\hat{\theta}^{\mathrm{p}} - \theta^{\mathrm{p}}\right)^2,$$

so truthful private-cost reporting is uniquely optimal with a quadratic margin.

Sequential verification. After allocation, evidence arrives sequentially as

$$X_l \sim_{iid} N(\theta^{\mathrm{e}}, 1), \qquad l = 1, ..., k.$$

For a reported environmental type $\hat{\theta}^{\mathrm{e}}$, we use the standard Gaussian normal-mixture e-process

$$E_l(\hat{\theta}^{\mathrm{e}}) = \frac{1}{\sqrt{1+l}} \exp\left\{ \frac{\left(\sum_{t=1}^{l}(X_t - \hat{\theta}^{\mathrm{e}})\right)^2}{2(1+l)} \right\},$$

following the normal-mixture construction used in anytime-valid inference; see Howard et al. [2021]. Verification terminates at the first threshold crossing,

$$\tau = \inf\left\{ l : E_l(\hat{\theta}^{\mathrm{e}}) \geq 1/\alpha \right\}.$$

The adjustable benchmark payment P^* is divided evenly across the k stages. If verification terminates at stage τ, the payment at that stage and all subsequent payments are withheld. Thus, the fraction of the benchmark payment retained is

$$\phi_k = \begin{cases} (\tau - 1)/k, & \tau \leq k \\ 1, & \tau > k. \end{cases}$$

Under calibrated correction, expected utility is therefore

$$\mathbb{E}\left[\hat{U}(\widehat{\theta}; \theta, X_k)\right] = \frac{\mathbb{E}\left[\phi_k\right]}{1-\alpha} P^*(\hat{\theta}^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}}) - \theta^{\mathrm{p}} q^*(\hat{\theta}^{\mathrm{p}}, \hat{\theta}^{\mathrm{e}}). \tag{13}$$

We compare the oracle benchmark and no verification with four verification schemes: fixed-level verification without uplift uses $\alpha = 0.1$ but does not apply the factor $1/(1-\alpha)$, calibrated corrections

$$\alpha_k^{\text{fast}} = e^{-\sqrt{k}}, \qquad \alpha_k^{\text{slow}} = k^{-1/4},$$

and the balanced choice defined by

$$\alpha_k^{\text{balanced}} = \sqrt{\frac{\log(1/\alpha_k^{\text{balanced}})}{k}}.$$

The fast sequence provides stronger protection against erroneous correction but raises the e-process threshold, which slows the detection of environmental misreports. The slow sequence yields sharper verification but retains a larger finite-sample calibration uplift. The balanced choice implements the rate-balancing prescription $\alpha_k \asymp s_k(\alpha_k)$.

We set $\theta^{\mathrm{p}} = 0.8$ and $\theta^{\mathrm{e}} = 0.5$. For each evidence horizon k, we estimate $\mathbb{E}\left[\phi_k\right]$ using common-random-number Monte Carlo paths and numerically compute the best-response environmental report; conditional on $\hat{\theta}^{\mathrm{e}}$, the optimal cost report is obtained analytically. The simulation uses 20,000 Monte Carlo paths and a grid of 1,000 environmental reports.

D Proof

D.1 Proof of Theorem 1

Fix an agent i, its type θ_i, and a profile of opponents' reports $\widehat{\theta}_{-i}$. Recall the expected utility under the calibrated correction is

$$\mathbb{E}\left[\hat{U}_i(\widehat{\theta}; \theta_i, X_k)\right] = U_i^-(\widehat{\theta}; \theta_i) + \frac{\mathbb{E}\left[\phi_{ik}(\widehat{\theta}, X_k)\right]}{1-\alpha_k} \left(U_i^*(\widehat{\theta}; \theta_i) - U_i^-(\widehat{\theta}; \theta_i) \right). \tag{14}$$

We start by showing that, under truthful reporting, the agent's expected utility under the calibrated correction is no worse than that of the truthful benchmark.

Lemma 5. *Under Assumption 1 and Assumption 2, for all sufficiently large k, agent i, type θ_i, and realized profile of opponents' reports $\widehat{\theta}_{-i}$,*

$$\mathbb{E}\left[\hat{U}_i(\theta_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] \geq U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i). \tag{15}$$

Since the benchmark mechanism is individually rational,

$$\mathbb{E}\left[\hat{U}_i(\theta_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] \geq U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) \geq 0.$$

Next, we bound the utility of an arbitrary deviation. We show below that, with the calibrated correction, the agent couldn't improve their utility by an amount larger than $\mathcal{O}(s_k + \alpha_k)$ over the truthful benchmark.

Lemma 6. *Under Assumption 1–Assumption 3, for all sufficiently large k, agent i, type θ_i, and realized profile of opponents' reports $\widehat{\theta}_{-i}$, there exists a constant $C < \infty$ such that*

$$\mathbb{E}\left[\hat{U}_i(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] \leq U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) + L_B C s_k + \frac{\alpha_k}{1-\alpha_k}\overline{\Delta}. \tag{16}$$

Combining Lemma 5 and Lemma 6,

$$\begin{aligned}
&\sup_{\widehat{\theta}_i} \mathbb{E}\left[\hat{U}_i(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] - \mathbb{E}\left[\hat{U}_i(\theta_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] \\
&\qquad \leq U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) + L_B C s_k + \frac{\alpha_k}{1-\alpha_k}\overline{\Delta} - U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) \\
&\qquad = \mathcal{O}(s_k + \alpha_k).
\end{aligned}$$

D.2 Proof of Proposition 2

Under the Clarke pivot payments, agent i's utility equals its marginal contribution to welfare, and thus it achieves a utility of

$$\begin{aligned}
B_i^{VCG}(e) = \sup_{\hat{\theta}_i^{\mathrm{p}}} &\left\{ \theta_i^{\mathrm{p}} q_i^{VCG}(\hat{\theta}_i^{\mathrm{p}}) + F\left(q^{VCG}(\hat{\theta}_i^{\mathrm{p}}), (e, \hat{\theta}_{-i}^{\mathrm{e}})\right) + \sum_{j \neq i} \hat{\theta}_j^{\mathrm{p}} q_j^{VCG}(\hat{\theta}_i^{\mathrm{p}}) \right\} \\
&- \max_{q \in \mathcal{Q}, q_i = 0} \left\{ F(q, (e, \hat{\theta}_{-i}^{\mathrm{e}})) + \sum_{j \neq i} \hat{\theta}_j^{\mathrm{p}} q_j \right\},
\end{aligned}$$

where the last term is the VCG externality term and we drop the dependence of q^{VCG} on the other arguments for notational simplicity. Since VCG is DSIC in $\hat{\theta}_i^{\mathrm{p}}$,

$$\begin{aligned}
B_i^{VCG}(e) &= \theta_i^{\mathrm{p}} q_i^{VCG}(\theta_i^{\mathrm{p}}) + F\left(q^{VCG}(\theta_i^{\mathrm{p}}), (e, \hat{\theta}_{-i}^{\mathrm{e}})\right) + \sum_{j \neq i} \hat{\theta}_j^{\mathrm{p}} q_j^{VCG}(\theta_i^{\mathrm{p}}) - \max_{q \in \mathcal{Q}, q_i = 0} \left\{ F(q, (e, \hat{\theta}_{-i}^{\mathrm{e}})) + \sum_{j \neq i} \hat{\theta}_j^{\mathrm{p}} q_j \right\} \\
&= \max_{q \in \mathcal{Q}} \left\{ F(q, (e, \hat{\theta}_{-i}^{\mathrm{e}})) + \theta_i^{\mathrm{p}} q_i + \sum_{j \neq i} \hat{\theta}_j^{\mathrm{p}} q_j \right\} - \max_{q \in \mathcal{Q}, q_i = 0} \left\{ F(q, (e, \hat{\theta}_{-i}^{\mathrm{e}})) + \sum_{j \neq i} \hat{\theta}_j^{\mathrm{p}} q_j \right\}.
\end{aligned}$$

Now, since a maximum of functions that are uniformly L_F-Lipschitz is itself L_F-Lipschitz, it follows directly from Equation (10) that

$$\left|B_i^{VCG}(e) - B_i^{VCG}(e')\right| \leq 2L_F \left\|e - e'\right\|.$$

D.3 Proof of Proposition 3

Define the optimal virtual-surplus value

$$H_i(p, e) := \max_{q \in \mathcal{Q}} \left\{ F(q, (e, \hat{\theta}_{-i}^{\mathrm{e}})) + \psi_i(p) q_i + \sum_{j \neq i} \psi_j(\hat{\theta}_j^{\mathrm{p}}) q_j \right\}.$$

Equation (10) immediately gives that

$$|H_i(p, e) - H_i(p, e')| \leq L_F \left\|e - e'\right\| \tag{17}$$

uniformly in p. Since $q_i \in [0,1]$ and ψ_i has bounded derivative on the compact type interval, $H_i(p,e)$ is also Lipschitz in p, and thus differentiable almost everywhere. At every differentiability point, Danskin's theorem gives

$$\frac{\partial H_i(p,e)}{\partial p} = \psi_i'(p) q_i^M\left((p, \hat{\theta}^{\mathrm{p}}_{-i}), (e, \hat{\theta}^{\mathrm{e}}_{-i})\right).$$

On the other hand, the Myerson payment identity and DSIC imply

$$B_i^M(e) = \int_{\underline{\theta^{\mathrm{p}}}}^{\theta_i^{\mathrm{p}}} \frac{1}{\psi_i'(p)} \frac{\partial H_i(p,e)}{\partial p} dp,$$

under the usual normalization that the lowest private type receives zero utility, and thus

$$\begin{aligned} B_i^M(e) - B_i^M(e') &= \int_{\underline{\theta^{\mathrm{p}}}}^{\theta_i^{\mathrm{p}}} \frac{\partial_p H_i(p,e) - \partial_p H_i(p,e')}{\psi_i'(p)} dp \\ &= \left[\frac{H_i(p,e) - H_i(p,e')}{\psi_i'(p)}\right]_{\underline{\theta^{\mathrm{p}}}}^{\theta_i^{\mathrm{p}}} \\ &\qquad - \int_{\underline{\theta^{\mathrm{p}}}}^{\theta_i^{\mathrm{p}}} \left(H_i(p,e) - H_i(p,e')\right) \frac{d}{dp}\left(\frac{1}{\psi_i'(p)}\right) dp. \end{aligned}$$

Since ψ_i is twice continuously differentiable and ψ_i' is uniformly bounded away from zero, it follows directly from Equation (17) that Proposition 3 holds.

D.4 Proof of Proposition 4

Fix an agent i, its type θ_i, a profile of opponents' reports $\widehat{\theta}_{-i}$, and an active exact best response $\widehat{\theta}_{ik} = (\hat{\theta}^{\mathrm{p}}_{ik}, \hat{\theta}^{\mathrm{e}}_{ik})$. Throughout, we omit the dependence of $\widehat{\theta}_{ik}$ on k for notational simplicity.

From Assumption 2, $\rho(t) \to 0$. Thus, we can choose $C < \infty$ sufficiently large such that $\overline{\Delta}\rho(C) < \kappa$. Since $\alpha_k, s_k \to 0$, for all sufficiently large k,

$$\frac{\overline{\Delta}\rho(C)}{1-\alpha_k} < \kappa, \qquad C s_k \le r_0.$$

First, we show that $\left\|\hat{\theta}^{\mathrm{e}}_i - \theta^{\mathrm{e}}_i\right\| = \mathcal{O}(s_k)$. Suppose for contradiction that $\left\|\hat{\theta}^{\mathrm{e}}_i - \theta^{\mathrm{e}}_i\right\| > C s_k$, then Assumption 1 and Assumption 2 imply that

$$\begin{aligned} \mathbb{E}\left[\hat{U}_i(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] &= U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) + \frac{\mathbb{E}\left[\phi_{ik}(\widehat{\theta}_i, \widehat{\theta}_{-i}, X_k)\right]}{1-\alpha_k}\left(U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i)\right) \\ &\le U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) + \frac{\rho(C)}{1-\alpha_k}\left(U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i)\right) \\ &\le U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) - \kappa + \frac{\overline{\Delta}\rho(C)}{1-\alpha_k} \\ &< U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i), \end{aligned}$$

By calibrated validity, truthful reporting gives expected utility at least $U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i)$, which contradicts the optimality of $\widehat{\theta}_i$. Thus, $\left\|\hat{\theta}^{\mathrm{e}}_i - \theta^{\mathrm{e}}_i\right\| \le C s_k$.

Next, we show that $\left\|\hat{\theta}^{\mathrm{p}}_i - \theta^{\mathrm{p}}_i\right\| = \mathcal{O}\left((s_k + \alpha_k)^{1/\beta}\right)$. Since $0 \le \phi_{ik} \le 1$,

$$\begin{aligned} \mathbb{E}\left[\hat{U}_i(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] &= U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) + \frac{\mathbb{E}\left[\phi_{ik}(\widehat{\theta}_i, \widehat{\theta}_{-i}, X_k)\right]}{1-\alpha_k}\left(U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i)\right) \\ &\le U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) + \frac{1}{1-\alpha_k}\left(U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i)\right) \\ &\le U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) + \frac{\alpha_k}{1-\alpha_k}\overline{\Delta}. \end{aligned}$$

By Assumption 4,

$$U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) \leq U_i^*((\theta_i^{\mathrm{p}}, \hat{\theta}_i^{\mathrm{e}}), \widehat{\theta}_{-i}; \theta_i) - C_M \left(\left\| \hat{\theta}_i^{\mathrm{p}} - \theta_i^{\mathrm{p}} \right\| \wedge r_0 \right)^{\beta}.$$

Moreover, Assumption 3 and $\left\| \hat{\theta}_i^{\mathrm{e}} - \theta_i^{\mathrm{e}} \right\| \leq C s_k$ imply that

$$\begin{aligned} U_i^*((\theta_i^{\mathrm{p}}, \hat{\theta}_i^{\mathrm{e}}), \widehat{\theta}_{-i}; \theta_i) &\leq B(\hat{\theta}_i^{\mathrm{e}}) \\ &\leq B(\theta_i^{\mathrm{e}}) + L_B \left\| \hat{\theta}_i^{\mathrm{e}} - \theta_i^{\mathrm{e}} \right\| \\ &\leq U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) + L_B C s_k. \end{aligned}$$

But $\mathbb{E}\left[\hat{U}_i(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i, X_k) \right] \geq U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i)$ because $\widehat{\theta}_i$ is a best response. Thus,

$$\begin{aligned} U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) &\leq \mathbb{E}\left[\hat{U}_i(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i, X_k) \right] \\ &\leq U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) + \frac{\alpha_k}{1-\alpha_k} \overline{\Delta} \\ &\leq U_i^*((\theta_i^{\mathrm{p}}, \hat{\theta}_i^{\mathrm{e}}), \widehat{\theta}_{-i}; \theta_i) - C_M \left(\left\| \hat{\theta}_i^{\mathrm{p}} - \theta_i^{\mathrm{p}} \right\| \wedge r_0 \right)^{\beta} + \frac{\alpha_k}{1-\alpha_k} \overline{\Delta} \\ &\leq U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) + L_B C s_k - C_M \left(\left\| \hat{\theta}_i^{\mathrm{p}} - \theta_i^{\mathrm{p}} \right\| \wedge r_0 \right)^{\beta} + \frac{\alpha_k}{1-\alpha_k} \overline{\Delta}, \end{aligned}$$

and

$$C_M \left(\left\| \hat{\theta}_i^{\mathrm{p}} - \theta_i^{\mathrm{p}} \right\| \wedge r_0 \right)^{\beta} \leq L_B C s_k + \frac{\alpha_k}{1-\alpha_k} \overline{\Delta}.$$

The right-hand side converges to zero, so for all sufficiently large k, $\left\| \hat{\theta}_i^{\mathrm{p}} - \theta_i^{\mathrm{p}} \right\| < r_0$. Therefore,

$$\left\| \hat{\theta}_i^{\mathrm{p}} - \theta_i^{\mathrm{p}} \right\| = \mathcal{O}\left((s_k + \alpha_k)^{1/\beta} \right).$$

D.5 Proof of Lemma 5

From Equation (14), for a truthful agent,

$$\mathbb{E}\left[\hat{U}_i(\theta_i, \widehat{\theta}_{-i}; \theta_i, X_k) \right] = U_i^-(\theta_i, \widehat{\theta}_{-i}; \theta_i) + \frac{\mathbb{E}\left[\phi_{ik}(\theta_i, \widehat{\theta}_{-i}, X_k) \right]}{1-\alpha_k} \left(U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\theta_i, \widehat{\theta}_{-i}; \theta_i) \right).$$

By Assumption 2,

$$\mathbb{E}\left[\phi_{ik}(\theta_i, \widehat{\theta}_{-i}, X_k) \right] \geq 1 - \alpha_k.$$

By Assumption 1, $U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\theta_i, \widehat{\theta}_{-i}; \theta_i) \geq 0$, and thus

$$\begin{aligned} \mathbb{E}\left[\hat{U}_i(\theta_i, \widehat{\theta}_{-i}; \theta_i, X_k) \right] &= U_i^-(\theta_i, \widehat{\theta}_{-i}; \theta_i) + \frac{\mathbb{E}\left[\phi_{ik}(\theta_i, \widehat{\theta}_{-i}, X_k) \right]}{1-\alpha_k} \left(U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\theta_i, \widehat{\theta}_{-i}; \theta_i) \right) \\ &\geq U_i^-(\theta_i, \widehat{\theta}_{-i}; \theta_i) + \left(U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\theta_i, \widehat{\theta}_{-i}; \theta_i) \right) \\ &= U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) \end{aligned}$$

D.6 Proof of Lemma 6

From Assumption 2, $\rho(t) \to 0$. Thus, we can choose a constant $C < \infty$ such that $\overline{\Delta}\rho(C) < \kappa$. Since $\alpha_k, s_k \to 0$, for all sufficiently large k,

$$\frac{\overline{\Delta}\rho(C)}{1-\alpha_k} \leq \kappa, \qquad C s_k \leq r_0.$$

We consider the following two cases.

Case 1: $\left\|\hat{\theta}_i^{\mathrm{e}} - \theta_i^{\mathrm{e}}\right\| > Cs_k$**.** By Assumption 2 and monotonicity of ρ,

$$\mathbb{E}_{\theta_i^{\mathrm{e}}}\left[\phi_{ik}(\widehat{\theta}, X_k)\right] \leq \rho\left(\left\|\hat{\theta}_i^{\mathrm{e}} - \theta_i^{\mathrm{e}}\right\| / s_k\right) \leq \rho(C).$$

By Assumption 1,

$$\begin{aligned}
\mathbb{E}\left[\hat{U}_i(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] &= U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) + \frac{\mathbb{E}\left[\phi_{ik}(\widehat{\theta}_i, \widehat{\theta}_{-i}, X_k)\right]}{1-\alpha_k}\left(U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i)\right) \\
&\leq U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) + \frac{\rho(C)}{1-\alpha_k}\left(U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i)\right) \\
&\leq U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) - \kappa + \frac{\overline{\Delta}\rho(C)}{1-\alpha_k} \\
&\leq U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i),
\end{aligned}$$

where the last inequality follows from the construction.

Case 2: $\left\|\hat{\theta}_i^{\mathrm{e}} - \theta_i^{\mathrm{e}}\right\| \leq Cs_k$**.** Since $\mathbb{E}_{\theta_i^{\mathrm{e}}}\left[\phi_{ik}(\widehat{\theta}, X_k)\right] \leq 1$, by Assumption 1

$$\begin{aligned}
\mathbb{E}\left[\hat{U}_i(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] &= U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) + \frac{\mathbb{E}\left[\phi_{ik}(\widehat{\theta}_i, \widehat{\theta}_{-i}, X_k)\right]}{1-\alpha_k}\left(U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i)\right) \\
&\leq U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) + \frac{1}{1-\alpha_k}\left(U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) - U_i^-(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i)\right) \\
&\leq U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i) + \frac{\alpha_k}{1-\alpha_k}\overline{\Delta}.
\end{aligned} \tag{18}$$

By definition of B_i,

$$B_i(\hat{\theta}_i^{\mathrm{e}}) \geq U_i^*(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i).$$

Since $Cs_k \leq r_0$, Assumption 3 implies that

$$\begin{aligned}
B_i(\hat{\theta}_i^{\mathrm{e}}) &\leq B_i(\theta_i^{\mathrm{e}}) + L_B\left\|\hat{\theta}_i^{\mathrm{e}} - \theta_i^{\mathrm{e}}\right\| \\
&\leq U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) + L_B C s_k
\end{aligned}$$

Combining with Equation (18),

$$\begin{aligned}
\mathbb{E}\left[\hat{U}_i(\widehat{\theta}_i, \widehat{\theta}_{-i}; \theta_i, X_k)\right] &\leq B_i(\hat{\theta}_i^{\mathrm{e}}) + \frac{\alpha_k}{1-\alpha_k}\overline{\Delta} \\
&\leq U_i^*(\theta_i, \widehat{\theta}_{-i}; \theta_i) + L_B C s_k + \frac{\alpha_k}{1-\alpha_k}\overline{\Delta}.
\end{aligned}$$

This proves the claimed result.