Learning in Policy Transparency Games: Wedge Structure and Adaptive Certification

Uyen Nguyen[1], **Khanh Nguyen**[2,3,4], **Hai Nguyen**[3], **Ngoc Vu**[2], **Hung Mai**[1,2], **Luong Doan**[2], **Trang Le**[1,2], **Trang Pham**[2], **Nhung Duong**[2], **Tuan Do**[4*]
[1]National Economics University, Vietnam
[2]N2TP Technology, Vietnam
[3]Nanyang Technological University, Singapore
[4]Phenikaa School of Information Technology, Phenikaa University, Vietnam

Abstract

As learning agents are increasingly deployed in strategic environments, designers must choose whether to deploy transparent (auditable, committing) or opaque (flexible, reactive) policies. We formalize this through *Policy Transparency Games* (PTGs): a two-stage model in which each agent first chooses transparency, then plays an underlying normal-form game. Our central theoretical contribution is a structural reduction: for two-player PTGs, all transparency incentives decompose into two per-player quantities, a leader wedge $L_i - N_i$ and a follower wedge $F_i - N_i$, whose signs characterize every equilibrium and cleanly separate selection-robust from selection-dependent conclusions. We then ask how a learning agent can discover this structure from interaction data. External-regret learners over the binary transparency action reach coarse correlated equilibrium, and our adaptive WEDGE-CERTIFY algorithm attains instance-dependent certificate complexity $\tilde{O}(\sigma^2(d_0^{-2} + d_{10}^{-2} + d_{01}^{-2}))$ with matching information-theoretic lower bounds. The wedge representation thus determines the learnability of endogenous algorithmic transparency.

1 Introduction

The delegation of strategic decision-making to algorithmic agents introduces a modeling choice absent from classical game theory: when strategies are implemented as algorithms, players must decide whether their policies are observable to others. This observability choice fundamentally alters the strategic environment as it determines whether a player can credibly commit to a policy or instead retain flexibility by reacting to others.

We consider a firm A which deploys automated pricing systems. If Firm A's algorithm is transparent (open-source or auditable), competitors can anticipate its prescribed action distribution. This transparency functions as a credible commitment device since A cannot deviate from its published policy, including any randomization it specifies. Conversely, if A decides to make its algorithm opaque, A retains strategic flexibility and can condition on the announcements of transparent players, but loses its commitment power. In the baseline model below, we abstract transparency to commitment to an observable mixed action; richer policy-function transparency is left for future work. This isolates one strategic mechanism rather than fully explaining real-world open- versus closed-model decisions, which also involve safety, regulation, intellectual property, development costs, reputation, and repeated interactions.

*Correspondence to: tuan.dongoc@phenikaa-uni.edu.vn

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

To formalize this tension, we introduce *Policy Transparency Games* (PTGs), a two-stage framework in which players first simultaneously choose between transparent and opaque policies, and subsequently play an underlying normal-form game subject to the induced commitment structure. In contrast to classical commitment and Stackelberg models, where the commitment structure is fixed exogenously, our framework endogenizes the choice of commitment itself through strategic control of policy observability. This two-stage formulation enables a systematic characterization of equilibrium transparency choices and their efficiency implications across games, providing new predictions for algorithmic competition settings. Beyond this game-theoretic characterization, our learning-theoretic contribution is to identify which equilibrium-relevant comparisons must be learned from noisy interaction, how samples should be allocated, and which statistical bottlenecks are unavoidable. This connects endogenous transparency to multi-agent learning and adaptive equilibrium certification.

Transparency affects two comparisons at once. A player who becomes transparent is a leader against an opaque opponent and avoids being a follower against a transparent opponent. The paper's main message is that these two comparisons are distinct and jointly determine transparency incentives. We call them the leader wedge and follower wedge. Mutual transparency and mutual opacity can yield the same Stage-2 payoff while having different Stage-1 stability, because deviations from the two profiles expose different wedges.

1.1 Related Work

Stackelberg and Leadership Games. The strategic value of commitment was articulated by Schelling (1980). Subsequently, the Stackelberg leadership model (Stackelberg, 1934) formalized commitment games; the optimal commitment was proved to be polynomial-time and computable via LP under SSE by Conitzer and Sandholm (2006). von Stengel and Zamir (2010) established existence results for leadership games with convex strategy sets, and Letchford et al. (2014) quantified the "value of commitment".

Endogenous Commitment and Timing. Caruana and Einav (2008) developed endogenous commitment through costly delays, where switching costs create a "commitment ladder". Our PTG framework differs by modeling commitment as a discrete binary choice which is made simultaneously. Hamilton and Slutsky (1990) studied endogenous timing in duopoly; PTGs restrict this to a binary transparency decision.

Translucent Players and Program Equilibrium. Halpern and Pass (2018) modeled "translucent" players with probabilistic strategy observability. Relatedly, Tennenholtz (2004) introduced program equilibrium where players submit readable computer programs; Oesterheld (2019) extended this to robust variants. Kalai et al. (2010) proved that voluntary conditional commitments can achieve any feasible individually rational payoff.

Information Design and Strategic Delegation. Bayesian persuasion (Kamenica and Gentzkow, 2011; Bergemann and Morris, 2019) studies how an informed sender designs signals before strategic interaction. Commitment with signaling under asymmetric information, including Li and Zhu (2023), similarly combines observability with commitment. Strategic delegation (Fershtman et al., 1991) and observable-contract models such as Sourbut et al. (2024) study how committing to an instrument changes downstream play. PTGs ask a complementary question: when observability itself is optional and chosen simultaneously by all players, which commitment structure is selected endogenously?

Evolutionary and Security Game Applications. In applied settings, Bianca Ogbo et al. (2025) studied institutional incentives in coordination with commitments; Vallier (2025) developed "strategic replicators" theory. Stackelberg security games (Tambe, 2012; Korzhyk et al., 2011) assumed defenders commit; PTGs make transparency itself strategic.

Industrial Organization and Complexity. From an industrial organization perspective, Bulow et al. (1985) distinguished strategic substitutes from complements, and Fudenberg and Tirole (1984) taxonomized commitment effects. Our Cournot vs. Bertrand analysis shows how that distinction corresponds to the sign of the leader and follower wedges. From a computational standpoint, Chen et al. (2009) and Daskalakis et al. (2009) established PPAD-completeness for Nash equilibria, while Etessami and Yannakakis (2010) showed FIXP-completeness for games with $n \geq 3$ players. Coniglio et al. (2017) established NP-hardness for pessimistic multi-follower Stackelberg equilibria, with algorithms provided by Basilico et al. (2017). Our computational discussion separates the binary transparency layer from the selected Stage-2 equilibrium computations.

Learning in Games. We use the external-regret-to-coarse-correlated-equilibrium principle (Cesa-Bianchi and Lugosi, 2006) and develop adaptive certification theorems for wedge signs and local equilibrium certificates. The proofs combine anytime confidence sequences with change-of-measure lower bounds tailored to the comparison graphs induced by PTGs.

1.2 Contributions

We study two-player Policy Transparency Games (PTGs) through a wedge representation. For a fixed simultaneous-profile-consistent Stage-2 selection protocol χ, let $N_i := U_i^{\chi}(0,0) = U_i^{\chi}(1,1)$ denote player i's simultaneous payoff, and let L_i and F_i denote her sole-leader and sole-follower payoffs, respectively. We define the leader and follower wedges by $\ell_i := L_i - N_i, \qquad f_i := F_i - N_i$. These two wedges per player are the organizing quantities of the paper.

Our first contribution is the transparency-wedge theorem (Theorem 1), which shows that every selected two-player PTG is fully characterized by the four wedges $(\ell_1, f_1, \ell_2, f_2)$. Their signs give necessary and sufficient conditions for all pure transparency equilibria, identify when transparency or opacity is dominant, and characterize interior mixed equilibria.

Second, we prove a selection-robustness theorem (Theorem 2). For each transparency profile t, let $X(t)$ be the set of admissible Stage-2 payoff vectors. Then t is a Stage-1 equilibrium for every admissible simultaneous-profile-consistent selection protocol if and only if each player's worst admissible payoff at t is at least her best admissible payoff from any unilateral transparency deviation. This separates robust transparency conclusions from selection-dependent ones.

Third, we connect the wedge structure to learning and certification. No-regret transparency learners approach coarse correlated equilibrium of the induced binary transparency game, while strict wedge margins yield finite mistake bounds (Theorem 4). We also propose WEDGE-CERTIFY (Theorem 3), which samples only the three informative profiles and identifies the wedge signs with instance-dependent certificate complexity$\widetilde{O}\big(\sigma^2(d_0^{-2} + d_{10}^{-2} + d_{01}^{-2})\big)$,where $d_{10} := \min\{|\ell_1|, |f_2|\}$, $d_{01} := \min\{|f_1|, |\ell_2|\}$, and $d_0 := \min\{d_{10}, d_{01}\}$. Matching lower bounds show that these bottlenecks are unavoidable.

Finally, we scope the computational and economic implications of the wedge representation. Exact equilibrium computation is polynomial for 2×2 PTGs, while augmented variants inherit hardness from embedded Stage-2 subproblems. The Cournot and differentiated-Bertrand examples illustrate how the wedge signs recover the classical distinction between strategic substitutes and strategic complements.

1.3 Paper Organization

Section 2 introduces Policy Transparency Games and the induced Stage-1 transparency game. Section 3 presents the wedge characterization and the selection-robustness result. Section 4 develops adaptive wedge certification, and Section 5 illustrates the wedge regimes through benchmark games and industrial-organization examples. Additional proofs, learning extensions, computational results, and boundary examples are deferred to the appendix.

2 Model and Preliminaries

2.1 Base Games

Let $G = (N, (S_i)_{i\in N}, (u_i)_{i\in N})$ be a finite normal-form game where $N = \{1, 2, \ldots, n\}$ is the set of players, S_i is player i's finite action set with $|S_i| = m_i$, and $u_i : \prod_{j\in N} S_j \to \mathbb{R}$ is player i's payoff function, and let $\Delta(S_i)$ denote the set of probability distributions over S_i. A mixed strategy for player i is $\sigma_i \in \Delta(S_i)$. We write $\sigma = (\sigma_1, \ldots, \sigma_n)$ for a strategy profile and σ_{-i} for the profile excluding player i.

Definition 1 (Best Response)**.** *Player i's best response correspondence is:*

$$BR_i(\sigma_{-i}) = \arg\max_{\sigma_i \in \Delta(S_i)} \sum_{s\in S} \left(\prod_{j\in N} \sigma_j(s_j) \right) u_i(s) \tag{1}$$

Definition 2 (Nash Equilibrium)**.** *A strategy profile σ^* is a Nash equilibrium if $\sigma_i^* \in BR_i(\sigma_{-i}^*)$ for all $i \in N$.*

We distinguish between the minimax value $\underline{v}_i = \max_{\sigma_i} \min_{\sigma_{-i}} u_i(\sigma_i, \sigma_{-i})$, representing what player i can guarantee against worst-case opponents, and the Stackelberg leader value $L_i = \max_{\sigma_i} \max_{\sigma_{-i} \in BR_{-i}(\sigma_i)} u_i(\sigma_i, \sigma_{-i})$, representing what player i achieves when committing first with opponents best-responding favorably. In general-sum games, $L_i \geq \underline{v}_i$, with equality in zero-sum games.

2.2 Policy Transparency Games

Definition 3 (Policy Transparency Game)**.** *Given a base game G, the associated Policy Transparency Game PTG(G) is a two-stage game. First, Stage 1 (Transparency Choice), each player i simultaneously chooses $t_i \in \{0, 1\}$ where $t_i = 1$ means player i is transparent (policy observable to all) and $t_i = 0$ means player i is opaque (policy unobservable). After that, Stage 2 (Base Game with Induced Structure), given transparency profile $t = (t_1, \ldots, t_n)$, players engage in game G with commitment structure determined by t.*

The key insight is that transparency creates commitment: a transparent player must announce their strategy before observing opaque players' choices, while opaque players can condition on transparent players' announcements.

Definition 4 (Induced Game Structure)**.** *Given transparency profile t, let $T = \{i : t_i = 1\}$ be the set of transparent players and $O = \{i : t_i = 0\}$ be the set of opaque players. The Stage 2 game proceeds as: (1) transparent players in T simultaneously announce strategies $(\sigma_i)_{i \in T}$; (2) opaque players in O observe these announcements and simultaneously choose strategies $(\sigma_j)_{j \in O}$. This induces a Stackelberg-like structure where transparent players are "leaders" and opaque players are "followers."*

Remark 1 (Equivalence of $(1, 1)$ and $(0, 0)$)**.** *Under mixed-action transparency, mutual transparency is equivalent to simultaneous play: when all players announce binding mixed strategies simultaneously, no player gains a first-mover advantage. Hence, in the two-player case, $\mathrm{Eq}_2(1, 1) = \mathrm{Eq}_2(0, 0) = NE(G)$.More generally, profiles with $T = N$ or $O = N$ are simultaneous-move profiles, and transparency creates strategic asymmetry only when both transparent and opaque players are present.*

This equivalence is a modeling convention for mixed-action transparency. Here, a transparent player commits to a mixed action in the base game. If transparency instead means revealing an executable program, a reaction function, or a conditional policy that can inspect other revealed policies, then mutual transparency can differ sharply from simultaneous play. Program-equilibrium and voluntary-commitment models study that richer object. In this case, the four-wedge reduction need not hold: the Stage-1 game remains binary, but its general incentive differences are

$$g_{i,t_{-i}} = U_i^{\chi}(1, t_{-i}) - U_i^{\chi}(0, t_{-i}),$$

which no longer necessarily reduce to the leader and follower wedges. In particular, $(1, 1)$ may cease to be statistically redundant. The present paper isolates the timing value of observable mixed-action commitment; Section 6 lists policy-function transparency as a separate extension of the baseline model.

Selection Protocol Convention. Stage 2 can have multiple equilibria. We therefore include equilibrium selection as part of the PTG instance. A **Stage-2 selection protocol** is a map $\chi : \{0, 1\}^n \to \bigcup_{t \in \{0,1\}^n} \mathrm{Eq}_2(t)$ such that $\chi(t) \in \mathrm{Eq}_2(t)$ for every transparency profile t. The induced payoff at t is $U_i^{\chi}(t) := u_i(\chi(t))$.

For asymmetric two-player profiles, our default Stage-2 protocol is optimistic Strong Stackelberg Equilibrium (SSE): followers break best-response ties in favor of the leader. This is an equilibrium-selection convention, not a behavioral assumption that followers must sacrifice their own payoffs, which means all selected responses remain follower-optimal. Optimistic SSE fixes the leader value L_i, but it may not uniquely determine the follower payoff F_{-i} when multiple leader-optimal commitments or commitment-response pairs exist. Hence χ includes a fixed deterministic tie-breaking rule over the leader-optimal SSE outcome set and returns a full payoff vector (L_i, F_{-i}) at each asymmetric profile. Alternatively, among leader-optimal outcomes, χ may favor the follower's payoff before

resolving remaining ties deterministically. This preserves the leader value but may change the follower payoff and hence the numerical wedges. For simultaneous profiles $(0,0)$ and $(1,1)$, χ selects the same Nash payoff vector by Remark 1. Results stated "under χ" are selection-dependent, whereas selection-robust results hold for every admissible such χ.

2.3 Equilibrium Concepts

Definition 5 (Stage 2 Equilibrium for Two-Player PTG)**.** *Given transparency profile t with sets T (transparent) and O (opaque), Stage 2 is an extensive-form game where transparent players announce strategies first, then opaque players respond. For opaque players observing announcements $\sigma_T = (\sigma_i)_{i\in T}$, the best-response is computed as $BR_j(\sigma_T, \sigma_{O\setminus\{j\}}) = \arg\max_{\sigma_j\in\Delta(S_j)} u_j(\sigma_T, \sigma_j, \sigma_{O\setminus\{j\}})$ For the two-player case with one leader, the transparent player solves $\sigma_i^* \in \arg\max_{\sigma_i} u_i(\sigma_i, \sigma_j^*(\sigma_i))$ where $\sigma_j^*(\sigma_i)$ is the SSE best response.*

Definition 6 (PTG Equilibrium under a Selection Protocol)**.** *Fix a Stage-2 selection protocol χ. Define the induced Stage-1 meta-game with action sets $\{0,1\}$ for each player and payoffs $U_i^\chi(t) = u_i(\chi(t))$. A PTG equilibrium under χ is a Stage-1 strategy profile $\tau^* = (\tau_i^*)_{i\in N}$, where $\tau_i^* \in \Delta(\{0,1\})$, that is a Nash equilibrium of this induced meta-game.*

The base PTG model has no direct payoff from the transparency decision: transparency affects payoffs only through the induced Stage-2 timing structure. In Appendix C.10 we also consider an explicitly marked augmented PTG in which additive transparency incentives or costs $\kappa_i(t_i)$ enter total payoffs as $u_i(\sigma) + \kappa_i(t_i)$. Those augmented results are not claims about the base model.

2.4 Standing Assumptions

Unless explicitly stated otherwise, all theorems in Sections 3 and 4 (and the deferred computational results of Appendix C.10) are stated under the following standing assumptions, abbreviated **(SA)** when invoked. Each subsequent theorem invokes only the subset it uses; relaxations or strengthenings are flagged at the theorem statement.

Assumption 1 (SA1: Two-player base game)**.** *G is a finite two-player normal-form game (Definition 1). Multi-player extensions are handled separately and explicitly noted (e.g., the three-player constant-sum example in Appendix A; the multi-follower hardness remarks in Appendix C.10.2).*

Assumption 2 (SA2: Simultaneous-profile-consistent selection protocol)**.** *A Stage-2 selection protocol χ is fixed. It is simultaneous-profile consistent $U_i^\chi(0,0) = U_i^\chi(1,1) =: N_i, i = 1,2$. At asymmetric profiles, χ selects a Stage-2 equilibrium of the induced leader-follower subgame and returns a full payoff vector (L_i, F_{-i}). The default is optimistic SSE with a fixed deterministic tie-breaking rule over leader-optimal outcomes; this rule resolves cases where the leader value L_i is unique but the follower payoff F_{-i} is not. Under the nondegeneracy condition of Theorem 6, the relevant outcome set reduces to a constant-size finite endpoint set. Selection-robust statements quantify over all admissible such χ.*

Assumption 3 (SA3: Wedge notation)**.** *The leader and follower payoffs L_i, F_i and the wedges $\ell_i := L_i - N_i$, $f_i := F_i - N_i$ are as in Section 1.2 and Theorem 1.*

Assumption 4 (SA4: Bounded payoffs and full-information transparency feedback (learning))**.** *For the no-regret learning theorem (Theorem 4, deferred to Appendix C.7), the selected Stage-2 payoff $U_i^\chi(t)$ lies in a bounded interval, normalized to $[0,1]$ without loss of generality, and after each round player i observes the entire transparency-action payoff vector against the opponent's realized action ($U_i^\chi(0, t_{-i}^r)$ and $U_i^\chi(1, t_{-i}^r)$). This is the full-information feedback model on the binary transparency action set. A bandit variant (observing only the realized $U_i^\chi(t^r)$) is discussed in Appendix C.7 but requires replacing Hedge with EXP3.*

Assumption 5 (SA5: Profile-pull noise model (certification))**.** *For Theorems 3 and 5, a profile pull of transparency profile t returns $Y_i(t) = U_i^\chi(t) + \xi_i(t)$, where, conditional on the past, $\xi_i(t)$ is mean-zero and σ^2-sub-Gaussian. Pulls of different profiles are independent given the history.*

Assumption 6 (SA6: Confidence parameter)**.** *Adaptive certification theorems are stated for $\delta \in (0, 1/4)$; the bound on δ is used only to apply the binary-relative-entropy lower bound $\mathrm{kl}(1-\delta, \delta) = \Omega(\log(1/\delta))$.*

Assumption 7 (SA7: Computational hypotheses)**.** *Algorithm 2 and Theorem 6 assume rational base-game payoffs and a Stage-2 selection protocol χ that is computable in time polynomial in the input bit-length on the constant-size equilibrium sets it is invoked on.*

3 Two-Player Characterization

We now present our main characterization results for two-player games with complete proofs.

3.1 The Transparency-Wedge Decomposition

The benchmark classes have simple commitment-wedge sign patterns. We now give the general two-player reduction. The theorem is stated as an algebraic identity for an arbitrary simultaneous-profile-consistent selection protocol χ: given any such χ, the four wedges $(\ell_1, f_1, \ell_2, f_2)$ determine the Stage-1 game and its equilibria. Some of the four sign regimes are not realizable under the *default* optimistic SSE protocol; Lemma 1 below records this restriction. Theorem 2 then identifies which conclusions do not depend on the selection at all.

Definition 7 (Canonical 2×2 Game)**.** *A 2×2 game G has payoff matrices: $U_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, U_2 = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ where a_{ij} and b_{ij} are Player 1's and Player 2's payoff, respectively, when Player 1 plays row i and Player 2 plays column j.*

Theorem 1 (Transparency-Wedge Theorem)**.** *Fix a two-player PTG and a Stage-2 selection protocol χ that is simultaneous-profile consistent: it selects the same simultaneous-game payoff vector at $(0,0)$ and $(1,1)$. Let $N_i := U_i^\chi(0,0) = U_i^\chi(1,1)$ be player i's selected simultaneous payoff. Let L_i be player i's payoff when i is the sole transparent leader, and let F_i be player i's payoff when i is the sole opaque follower. Define the leader and follower wedges $\ell_i := L_i - N_i$, $f_i := F_i - N_i$. Then the Stage-1 game is completely determined by the four wedges:*

	$t_2 = 0$	$t_2 = 1$
$t_1 = 0$	(N_1, N_2)	$(N_1 + f_1, N_2 + \ell_2)$
$t_1 = 1$	$(N_1 + \ell_1, N_2 + f_2)$	(N_1, N_2)

The signs of the two wedges summarize player i's transparency incentives. If $\ell_i > 0$ and $f_i < 0$, transparency strictly dominates opacity; if $\ell_i < 0$ and $f_i > 0$, opacity strictly dominates transparency. If both wedges are positive, $\ell_i > 0$ and $f_i > 0$, player i prefers to choose the opposite transparency action from the opponent, whereas if both wedges are negative, $\ell_i < 0$ and $f_i < 0$, player i prefers to match the opponent's transparency action.

The pure Stage-1 Nash equilibria are exactly characterized by the corresponding unilateral-deviation inequalities: $(0,0) \in NE_1$ iff $\ell_1 \le 0$ and $\ell_2 \le 0$; $(1,1) \in NE_1$ iff $f_1 \le 0$ and $f_2 \le 0$; $(1,0) \in NE_1$ iff $\ell_1 \ge 0$ and $f_2 \ge 0$; and $(0,1) \in NE_1$ iff $f_1 \ge 0$ and $\ell_2 \ge 0$. The fully mixed Stage-1 equilibria and degenerate cases are characterized in Appendix B.1.

Proof Sketch. Under (SA2), the simultaneous profiles $(0,0)$ and $(1,1)$ have the same selected payoff (N_1, N_2), while the asymmetric profiles are determined by the leader and follower payoffs: $(1,0)$ gives $(N_1+\ell_1, N_2+f_2)$ and $(0,1)$ gives $(N_1+f_1, N_2+\ell_2)$. This yields the displayed Stage-1 matrix. For player 1, the payoff difference between transparency and opacity is $D_1(q) = (1-q)\ell_1 - qf_1$, where $q = \Pr[t_2 = 1]$; similarly, $D_2(p) = (1-p)\ell_2 - pf_2$ for player 2. The endpoint signs of these affine functions give the wedge-sign regimes and the pure-equilibrium conditions, while interior mixed equilibria are exactly the solutions of $D_1(q) = D_2(p) = 0$, i.e., $(Z_2 \cap (0,1)) \times (Z_1 \cap (0,1))$. The nondegenerate, degenerate, and boundary cases follow from the corresponding form of the indifference sets Z_i. The full proof is deferred to Appendix B.1. □

Theorem 2 (Selection-Robust Transparency Equilibria)**.** *For each transparency profile t, let $X(t) \subseteq \mathbb{R}^2$ be the set of payoff vectors generated by all admissible Stage-2 equilibria under t. Admissible selection protocols are required to obey the simultaneous-profile consistency convention: the selected simultaneous payoff at $(0,0)$ and $(1,1)$ is the same payoff vector. For player i, we define the worst and best admissible Stage-2 payoffs for player i at transparency profile t as $\underline{U}_i(t) := \inf_{x \in X(t)} x_i$,*

and $\overline{U}_i(t) := \sup_{x \in X(t)} x_i$, respectively. A pure transparency profile t is a Stage-1 Nash equilibrium for every admissible Stage-2 selection protocol if and only if, for every player i and the unique unilateral deviation $t'_i \neq t_i$, $\underline{U}_i(t) \geq \overline{U}_i(t'_i, t_{-i})$.

Proof Sketch. Sufficiency. For any admissible χ obeying (SA2), player i's selected payoff at t is bounded below by $\underline{U}_i(t)$ and the selected payoff after a unilateral deviation is bounded above by $\overline{U}_i(t'_i, t_{-i})$, so the assumed interval inequality rules out a profitable deviation under every admissible χ. *Necessity.* If the inequality fails for some player i, choose admissible payoff vectors $x \in X(t)$ and $y \in X(t'_i, t_{-i})$ with $x_i < y_i$ and define an admissible χ that selects x, y at the two relevant profiles and arbitrary admissible payoffs elsewhere; (SA2) is preserved because unilateral transparency deviations never connect $(0,0)$ directly to $(1,1)$, so the simultaneous-profile-consistency constraint imposes no restriction on the constructed deviation. Under this χ, player i has a strict profitable Stage-1 deviation from t. The full proof is deferred to Appendix B.2. □

Theorem 2 turns the selection rule from an ambiguity into a diagnostic. If the interval inequalities hold, the transparency conclusion is selection-free. If they fail, the Stage-2 selection rule is part of the strategic environment and must be specified explicitly.

4 Learning Transparency Choices

We now add a learning layer: players repeatedly choose whether to be transparent and update the transparency choice by an online learning rule. The learning model treats the selected Stage-2 payoff vector $U^\chi(t)$ as the payoff of the induced binary-action transparency game.

Feedback model (full-information transparency feedback). At the end of each round r, after the realized transparency profile $t^r = (t_i^r, t_{-i}^r)$ has been played and Stage-2 has been resolved by χ, player i observes the entire transparency-action payoff vector against the opponent's realized action: the two counterfactual payoffs $U_i^\chi(0, t_{-i}^r)$ and $U_i^\chi(1, t_{-i}^r)$. This is full-information feedback over the binary transparency action set (two scalars per round) rather than a bandit observation of only the realized payoff. It is appropriate for settings where the regulator, platform, or simulator can evaluate Stage-2 under both transparency choices for player i ex post (for example, by re-running the selected Stage-2 protocol with the same opponent profile t_{-i}^r). A bandit-feedback variant, where the learner observes only $U_i^\chi(t^r)$, is discussed in Appendix C.7 and requires replacing Hedge with a two-action bandit algorithm (e.g. EXP3), weakening the regret rate by the standard bandit factor.

4.1 Adaptive Wedge Certification from Partial Profile Feedback

The wedge decomposition gives a profile-efficient certification problem. To classify the Stage-1 equilibrium structure, it suffices to estimate three selected payoff vectors. Under simultaneous-profile consistency, all four signs are determined by the profiles $(0,0)$, $(1,0)$, and $(0,1)$:

$$\begin{aligned} \ell_1 &= U_1^\chi(1,0) - U_1^\chi(0,0), \quad & f_2 &= U_2^\chi(1,0) - U_2^\chi(0,0), \\ f_1 &= U_1^\chi(0,1) - U_1^\chi(0,0), \quad & \ell_2 &= U_2^\chi(0,1) - U_2^\chi(0,0). \end{aligned} \tag{2}$$

The profile $(1,1)$ is redundant for classification. The shared reference profile $(0,0)$ creates the central allocation problem: every wedge uses the same simultaneous payoff, while the two asymmetric profiles each certify two signs. The learning result is specific to the PTG comparison graph: the structural equivalence of mutual transparency and mutual opacity removes one profile from the experiment and makes the shared simultaneous payoff the bottleneck for all wedge signs.

Suppose that a profile pull of transparency profile t returns both players' selected payoffs with noise $Y_i(t) = U_i^\chi(t) + \xi_i(t)$, where, conditional on the past, $\xi_i(t)$ is mean-zero and σ^2-sub-Gaussian. We define the asymmetric-profile bottleneck gaps as $d_{10} := \min\{|\ell_1|, |f_2|\}$, $d_{01} := \min\{|f_1|, |\ell_2|\}$, $d_0 := \min\{d_{10}, d_{01}\}$.

Feedback and sample allocation. Unlike the full-information feedback used for repeated-play learning, WEDGE-CERTIFY requires no counterfactual-profile observations. Each oracle pull selects one realized profile and observes both players' noisy payoffs. Thus, if the certifier can choose profiles and pool these two coordinates, profile-level bandit feedback does not change the allocation across

$(0,0)$, $(1,0)$, and $(0,1)$. Without pooling, however, individual learners may lack coordinates needed to reconstruct all wedges. If profiles are generated endogenously rather than selected by the certifier, forced exploration or importance weighting may be required, with additional dependence on visitation probabilities and estimator variance. These decentralized settings are outside the stated certification guarantee.

Algorithm 1 WEDGE-CERTIFY

Input: confidence level δ, profile-pull oracle for $(0,0), (1,0), (0,1)$.
State: sample counts N_t, empirical payoff means $\widehat{U}_i(t)$, and active wedge set $A = \{\ell_1, f_2, f_1, \ell_2\}$.
Initialization: pull each of $(0,0), (1,0), (0,1)$ once and initialize the empirical means.
while $A \neq \emptyset$ **do**
    pull $(0,0)$;
    pull $(1,0)$ if $\ell_1 \in A$ or $f_2 \in A$;
    pull $(0,1)$ if $f_1 \in A$ or $\ell_2 \in A$;
    for every active wedge $w = U_i^{\chi}(a) - U_i^{\chi}(0,0)$, with $a \in \{(1,0),(0,1)\}$ **do**
        form the empirical interval

$$I_w = \widehat{U}_i(a) - \widehat{U}_i(0,0) \;+\; [-r(N_a) - r(N_{00}),\, r(N_a) + r(N_{00})]\,;$$

        if $0 \notin I_w$ **then**
            remove w from A and record its sign.
        end if
    end for
end while
Output: the four wedge signs and the corresponding strict pure-equilibrium classification from Theorem 1.

Sample complexity is counted in profile pulls. One loop performs up to three profile pulls. For a profile sampled $m \geq 1$ times, use the dyadic stitched confidence radius

$$r(m) := 2\sigma \sqrt{\frac{\log(288(1 + \lfloor \log_2 m \rfloor)^2/\delta)}{m}}. \tag{3}$$

Theorem 3 (Adaptive wedge certification, instance-optimal up to anytime factors)**.** *Fix $\delta \in (0, 1/4)$ and a two-player PTG under a simultaneous-profile-consistent selection protocol χ. Assume all four wedges are nonzero and payoff feedback satisfies the conditional sub-Gaussian profile-pull model above. Then WEDGE-CERTIFY returns the correct wedge signs, and hence the correct strict pure-equilibrium classification, with probability at least $1 - \delta$. Moreover, on this event, for each $a \in \{00, 10, 01\}$, we have $N_a = O\Big(\frac{\sigma^2}{d_a^2}\Big[\log\frac{1}{\delta} + \log\Big(1 + \log\Big(1 + \frac{\sigma^2}{d_a^2}\Big)\Big)\Big]\Big)$, where $d_{00} := d_0$. Equivalently, the total certificate complexity is $\widetilde{O}\big(\sigma^2(d_0^{-2} + d_{10}^{-2} + d_{01}^{-2})\big)$.*

Conversely, let consider the Gaussian selected-payoff-table oracle model: the instance is identified with its selected-payoff table $\theta = (U_i^{\chi}(t))_{i,t}$ subject to the simultaneous-profile-consistency constraint $U_i^{\chi}(0,0) = U_i^{\chi}(1,1)$, the algorithm adaptively samples any of the four profiles, and a sample at profile t returns each player's payoff plus independent $\mathcal{N}(0,\sigma^2)$ noise. The class of admissible payoff tables is the full set of vectors in $\mathbb{R}^8$ satisfying simultaneous-profile consistency; the converse is an oracle lower bound on this payoff-table class and is not asserted as a realization lower bound over finite base games G paired with admissible χ unless the perturbed payoff table can be realized by some (G,χ) pair. Let $N_{\mathrm{sim}} := N_{00} + N_{11}$ be the total number of simultaneous-profile pulls. Any δ-correct algorithm identifying the wedge-sign vector on this oracle class must satisfy

$$\mathbb{E}[N_{\mathrm{sim}}] = \Omega\Big(\frac{\sigma^2}{d_0^2}\log\frac{1}{\delta}\Big), \quad \mathbb{E}[N_{10}] = \Omega\Big(\frac{\sigma^2}{d_{10}^2}\log\frac{1}{\delta}\Big), \quad \mathbb{E}[N_{01}] = \Omega\Big(\frac{\sigma^2}{d_{01}^2}\log\frac{1}{\delta}\Big). \tag{4}$$

Thus the wedge decomposition yields a certificate complexity $\Theta\big(\sigma^2\left(d_0^{-2} + d_{10}^{-2} + d_{01}^{-2}\right)\log(1/\delta)\big)$ up to the anytime logarithmic factor incurred by the stated algorithm. This improves over uniform estimation of all transparency profiles and proves that the improvement is information-theoretically limited.

Realizability of the oracle lower bound. The converse need not extend to every class of realizable finite games. Under optimistic SSE, for example, the leader wedge satisfies $\ell_i = L_i - N_i \geq 0$, whereas the oracle payoff-table class allows $\ell_i < 0$. Thus a perturbation that flips a wedge sign need not correspond to an admissible base game. A realization-class lower bound would require sign-changing alternatives that remain feasible under the chosen Stage-2 protocol. When such local alternatives exist, the same inverse-squared-gap scaling may follow; otherwise, the relevant complexity may depend on the distance to the nearest realizable sign-changing alternative. Characterizing this geometry remains open.

Uniform estimation of all four transparency profiles spends samples on $(1, 1)$, whereas the lower bound charges only the combined simultaneous representative $N_{\text{sim}} = N_{00} + N_{11}$. The gain is the removal of the redundant mutual-transparency profile and the optimal allocation of samples across the wedge comparison graph, not a generic concentration improvement.

Proof Sketch. Under (SA5)–(SA6), the dyadic stitched confidence sequence gives $|\widehat{U}_i^m(t) - U_i^X(t)| \leq r(m)$ uniformly over all players, informative profiles $t \in \{(0,0), (1,0), (0,1)\}$, and times $m \geq 1$, with probability at least $1 - \delta$. Hence each wedge interval contains its true value, and whenever an interval excludes zero, its sign is certified correctly; Theorem 1 then maps the certified signs to the strict pure-equilibrium classification. Inverting the confidence radius at the bottleneck gaps d_0, d_{10}, d_{01} gives the stated sample bounds. The lower bound follows from the adaptive transcript-KL identity in the Gaussian payoff-table oracle model: flipping the smallest incident wedge at each informative profile and applying the binary-testing lower bound $\mathrm{kl}(1 - \delta, \delta) = \Omega(\log(1/\delta))$ yields the matching profile-level lower bounds. The full proof is deferred to Appendix B.3. □

Theorem 3 is the learning-theoretic payoff of the wedge representation. The proof uses the geometry of the four wedge comparisons, the redundant $(1, 1)$ profile, and the shared simultaneous reference profile to obtain the adaptive certificate complexity.

5 Examples and Scope

This section illustrates the wedge mechanism in benchmark games and standard industrial-organization models. The Cournot/Bertrand contrast is the central punchline: the same wedge calculus records both the strategic-substitutes and strategic-complements cases. The remaining examples calibrate the wedge representation on classical benchmarks and mark the boundary of the two-player reduction.

Cournot. In symmetric interior linear Cournot quantity competition, the familiar Stackelberg ordering is $\pi^L > \pi^{NE} > \pi^F$. Therefore $\ell_i = \pi^L - \pi^{NE} > 0$ and $f_i = \pi^F - \pi^{NE} < 0$. Transparency strictly dominates opacity for both players, and the Stage-1 equilibrium is mutual transparency $(1, 1)$. This is the "strategic substitutes" case: commitment as a leader helps, while being the follower hurts.

Differentiated Bertrand. In the standard differentiated-price model with strategic complements and interior demands, the Stackelberg ordering is $\pi^F > \pi^L > \pi^{NE}$. Thus both wedges are positive: $\ell_i > 0$ and $f_i > 0$. Players prefer to have the opposite transparency status from the opponent, so the Stage-1 equilibria are asymmetric, $(1, 0)$ and $(0, 1)$. The contrast with Cournot illustrates how the wedge calculus records the strategic-complements case.

Constant-sum and coordination. In two-player constant-sum games the minimax theorem gives $L_i = F_i = N_i$ for both players, so all wedges vanish and the Stage-1 transparency game is payoff-trivial. With any positive transparency cost, mutual opacity becomes unique. In pure coordination games, if simultaneous subgames select a payoff-dominant equilibrium, then $L_i = F_i = N_i = v^*$ and all wedges again vanish; without that selection convention, transparency conclusions are selection-dependent, exactly as Theorem 2 indicates.

Battle of the Sexes. With mixed simultaneous-equilibrium selection, both the leader and follower payoffs in the asymmetric transparency profiles exceed the simultaneous mixed payoff. Thus $\ell_i > 0$ and $f_i > 0$ for both players. The wedge theorem immediately gives two asymmetric transparency equilibria, $(1, 0)$ and $(0, 1)$, without a separate case analysis. This places Battle of the Sexes in the same wedge regime as differentiated Bertrand. The analytical certification example in Appendix C.9 further illustrates how WEDGE-CERTIFY allocates samples across the three informative profiles and

how certifying a single asymmetric equilibrium can avoid one profile. It illustrates the theoretical prediction rather than providing empirical stopping-time validation.

Beyond two players. The two-player constant-sum invariance does not extend automatically to $n \geq 3$, because the opaque followers no longer act as a single minimizer. More generally, each player faces up to 2^{n-1} opponent-transparency configurations, with incentives $\Delta_i(t_{-i}) = U_i^X(1, t_{-i}) - U_i^X(0, t_{-i})$ that can depend on the identities and composition of transparent and opaque opponents. These differences need not collapse to two role-specific wedges. There are three-player constant-sum games in which a player improves by committing first; Appendix A gives an explicit construction. The counterexample has strictly dominant follower responses, so even uniqueness of the follower equilibrium does not restore the two-player invariance. Potential extensions may exploit coordinated followers, anonymous interactions, or separable payoff dependencies. This marks a boundary of the two-player wedge reduction.

6 Conclusion

We introduced Policy Transparency Games as a framework for analyzing and learning the strategic choice of commitment through observable policies. The main structural result is the transparency-wedge decomposition: any selected two-player PTG is summarized by the leader and follower wedges $\ell_i = L_i - N_i$ and $f_i = F_i - N_i$. Their signs determine whether transparency is dominant, opacity is dominant, or players coordinate/anti-coordinate over transparency choices, turning the two-stage model into a tractable diagnostic for endogenous commitment.

The selection-robustness theorem separates conclusions that hold under all admissible Stage-2 selection protocols from those that depend on equilibrium selection. The learning results show that the wedge structure is statistically discoverable: WEDGE-CERTIFY samples only the informative profiles, exploits the shared simultaneous reference payoff, and achieves instance-dependent certificate complexity $\tilde{O}(\sigma^2(d_0^{-2} + d_{10}^{-2} + d_{01}^{-2}))$, with matching lower bounds up to logarithmic factors. These results connect the static wedge characterization to algorithmic agents that learn transparency incentives from interaction data.

The examples show that the same wedge logic recovers standard economic distinctions: Cournot competition makes transparency dominant, while differentiated Bertrand competition generates asymmetric transparency equilibria. The Battle of the Sexes example also illustrates the predicted certification allocation, although finite-sample empirical validation remains future work. The baseline model treats transparency as commitment to a mixed action, so future work should study richer policy-function transparency, stochastic or continuous transparency, incomplete information, and nonstationary Stage-2 learning. Other open directions include structured multi-player extensions, realization-specific certification lower bounds, and decentralized certification under restricted feedback.

Funding Statement & Competing Interests

The authors received no external funding specifically, and there are no competing financial interests related to this research.

References

Basilico, N., Coniglio, S., Gatti, N., and Marchesi, A. (2017). Bilevel Programming Approaches to the Computation of Optimistic and Pessimistic Single-Leader-Multi-Follower Equilibria. In Iliopoulos, C. S., Pissis, S. P., Puglisi, S. J., and Raman, R., editors, *16th International Symposium on Experimental Algorithms (SEA 2017)*, volume 75 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 31:1–31:14, Dagstuhl, Germany. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

Bergemann, D. and Morris, S. (2019). Information Design: A Unified Perspective. *Journal of Economic Literature*, 57(1):44–95.

Bianca Ogbo, N., Song, Z., and Han, T. A. (2025). Evolution of Coordination Through Institutional Incentives: An Evolutionary Game Theory Approach. In *Proceedings of the 2025 7th International*

Conference on Distributed Artificial Intelligence, DAI '25, pages 38–47, New York, NY, USA. Association for Computing Machinery.

Bulow, J. I., Geanakoplos, J. D., and Klemperer, P. D. (1985). Multimarket Oligopoly: Strategic Substitutes and Complements. *Journal of Political Economy*, 93(3):488–511.

Caruana, G. and Einav, L. (2008). A Theory of Endogenous Commitment. *The Review of Economic Studies*, 75(1):99–116.

Cesa-Bianchi, N. and Lugosi, G. (2006). *Prediction, Learning, and Games*. Cambridge University Press, Cambridge.

Chen, X., Deng, X., and Teng, S.-H. (2009). Settling the complexity of computing two-player Nash equilibria. *J. ACM*, 56(3).

Coniglio, S., Gatti, N., and Marchesi, A. (2017). Pessimistic leader-follower equilibria with multiple followers. In Carles Sierra, I.-C., editor, *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, pages 171–177. International Joint Conferences on Artificial Intelligence Organization. Main track.

Conitzer, V. and Sandholm, T. (2006). Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, EC '06, pages 82–90, New York, NY, USA. Association for Computing Machinery.

Daskalakis, C., Goldberg, P. W., and Papadimitriou, C. H. (2009). The complexity of computing a Nash equilibrium. *Commun. ACM*, 52(2):89–97.

Etessami, K. and Yannakakis, M. (2010). On the Complexity of Nash Equilibria and Other Fixed Points. *SIAM Journal on Computing*, 39(6):2531–2597.

Fershtman, C., Judd, K. L., and Kalai, E. (1991). Observable Contracts: Strategic Delegation and Cooperation. *International Economic Review*, 32(3):551–559.

Fudenberg, D. and Tirole, J. (1984). The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look. *The American Economic Review*, 74(2):361–366.

Halpern, J. Y. and Pass, R. (2018). Game theory with translucent players. *International Journal of Game Theory*, 47(3):949–976.

Hamilton, J. H. and Slutsky, S. M. (1990). Endogenous timing in duopoly games: Stackelberg or cournot equilibria. *Games and Economic Behavior*, 2(1):29–46.

Kalai, A. T., Kalai, E., Lehrer, E., and Samet, D. (2010). A commitment folk theorem. *Games and Economic Behavior*, 69(1):127–137.

Kamenica, E. and Gentzkow, M. (2011). Bayesian Persuasion. *The American Economic Review*, 101(6):2590–2615.

Korzhyk, D., Yin, Z., Kiekintveld, C., Conitzer, V., and Tambe, M. (2011). Stackelberg vs. Nash in Security Games: An Extended Investigation of Interchangeability, Equivalence, and Uniqueness. *Journal of Artificial Intelligence Research*, 41:297–327.

Letchford, J., Korzhyk, D., and Conitzer, V. (2014). On the value of commitment. *Autonomous Agents and Multi-Agent Systems*, 28(6):986–1016.

Li, T. and Zhu, Q. (2023). Commitment with Signaling under Double-sided Information Asymmetry.

Oesterheld, C. (2019). Robust program equilibrium. *Theory and Decision*, 86(1):143–159.

Schelling, T. C. (1980). *The Strategy of Conflict: With a New Preface by the Author*. Harvard University Press.

Sourbut, O., Hammond, L., and Wood, H. (2024). Cooperation and control in delegation games. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, IJCAI '24.

Stackelberg, H. v. (1934). *Marktform und Gleichgewicht*. J. Springer.

Tambe, M. (2012). *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge University Press.

Tennenholtz, M. (2004). Program equilibrium. *Games and Economic Behavior*, 49(2):363–373.

Vallier, K. (2025). The Theory of Strategic Evolution: Games with Endogenous Players and Strategic Replicators.

von Stengel, B. and Zamir, S. (2010). Leadership games with convex strategy sets. *Games and Economic Behavior*, 69(2):446–457.

A A Three-Player Constant-Sum Boundary Example

The following zero-sum game shows that the two-player constant-sum invariance need not extend to three players. Player 1 chooses A or B, player 2 chooses L or R, and player 3 chooses U or D. Each entry lists (u_1, u_2, u_3), and every entry sums to zero.

If player 1 chooses A:

	U	D
L	$(1, -2, 1)$	$(-1, 1, 0)$
R	$(-1, -3, 4)$	$(-1, 0, 1)$

If player 1 chooses B:

	U	D
L	$(2, -1, -1)$	$(0, -2, 2)$
R	$(0, 0, 0)$	$(0, -1, 1)$

In the simultaneous game, (B, R, D) is the unique pure Nash equilibrium. At (B, R, D), player 1 receives 0 and receives -1 after deviating to A against (R, D); given B, player 2 prefers R to L when player 3 chooses D, and player 3 prefers D to U when player 2 chooses R. The other pure profiles fail at least one best-response condition: after A, the followers' unique best-response profile is (L, U), but player 1 then prefers B; after B, the followers' unique best-response profile is (R, D).

If player 1 is the sole transparent leader and commits to A, then in the follower subgame player 2 has a strictly dominant action L and player 3 has a strictly dominant action U. The unique follower equilibrium is therefore (L, U), giving player 1 payoff 1. Thus player 1 can improve from the simultaneous payoff 0 to the commitment payoff 1 even though the game is constant-sum.

B Full Proofs

This appendix collects the full proofs of the main theorems, deferred from the main text. All proofs invoke the standing assumptions **(SA1)-(SA7)** of Section 2.4, with each proof citing only the items it uses.

B.1 Proof of Theorem 1 (Transparency-Wedge)

For the fully mixed Stage-1 equilibria, we define the affine player-difference functions as $D_1(q) = (1-q)\ell_1 - qf_1, D_2(p) = (1-p)\ell_2 - pf_2$. A profile $(p, q) \in (0, 1)^2$ is a fully mixed Stage-1 equilibrium iff $D_1(q) = D_2(p) = 0$. Equivalently, for each player i the opponent's mixing probability must lie in player i's indifference set

$$Z_i = \begin{cases} \left\{ \dfrac{\ell_i}{\ell_i + f_i} \right\}, & \ell_i + f_i \neq 0, \\ [0, 1], & \ell_i = f_i = 0, \\ \varnothing, & \ell_i + f_i = 0 \text{ and } \ell_i \neq 0. \end{cases} \tag{5}$$

The set of fully mixed Stage-1 equilibria is therefore $\big(Z_2 \cap (0, 1)\big) \times \big(Z_1 \cap (0, 1)\big)$. In the nondegenerate case $\ell_i + f_i \neq 0$ for both players, Z_i is a singleton; the fully mixed equilibrium is then unique, exists iff both probabilities lie in $(0, 1)$, and is given by $p^* = \Pr[t_1 = 1] = \frac{\ell_2}{\ell_2 + f_2}$, and $q^* = \Pr[t_2 = 1] = \frac{\ell_1}{\ell_1 + f_1}$. In the fully degenerate case $\ell_1 = f_1 = \ell_2 = f_2 = 0$ (e.g., the two-player constant-sum benchmark), every profile in $(0, 1)^2$ is a fully mixed equilibrium and the set of equilibria is a continuum. Boundary and partial-degeneracy cases reduce to the pure-equilibrium inequalities above together with the appropriate Z_i restriction.

We invoke (SA1, SA2, SA3).

Payoff matrix. At the simultaneous profiles $(0, 0)$ and $(1, 1)$, (SA2) gives $U_i^{\chi}(0, 0) = U_i^{\chi}(1, 1) = N_i$ for both players. At the asymmetric profile $(1, 0)$, player 1 is the sole transparent leader and player 2 is the sole opaque follower, so the asymmetric-profile selection of χ delivers payoffs $L_1 = N_1 + \ell_1$

to player 1 and $F_2 = N_2 + f_2$ to player 2; the theorem makes no use of which Stage-2 selection rule is used to define L_i, F_i, only that some such rule is fixed. The profile $(0, 1)$ is symmetric. This proves the displayed Stage-1 payoff matrix.

Sign cases for player 1. Fix player 1 and let $q = \Pr[t_2 = 1]$. If player 1 chooses transparency ($t_1 = 1$), her expected payoff is

$$qU_1^X(1,1) + (1-q)U_1^X(1,0) = qN_1 + (1-q)(N_1 + \ell_1) = N_1 + (1-q)\ell_1.$$

If she chooses opacity ($t_1 = 0$), her expected payoff is

$$qU_1^X(0,1) + (1-q)U_1^X(0,0) = q(N_1 + f_1) + (1-q)N_1 = N_1 + qf_1.$$

Hence player 1's expected payoff difference between transparency and opacity is the affine function

$$D_1(q) = (1-q)\ell_1 - qf_1,$$

with $D_1(0) = \ell_1$ and $D_1(1) = -f_1$. The four wedge-sign cases (1)-(4) follow from the signs of $D_1(0)$ and $D_1(1)$:

- (1) $\ell_1 > 0, f_1 < 0$: $D_1 > 0$ on $[0, 1]$, so transparency strictly dominates opacity.
- (2) $\ell_1 < 0, f_1 > 0$: $D_1 < 0$ on $[0, 1]$, so opacity strictly dominates.
- (3) $\ell_1 > 0, f_1 > 0$: D_1 decreases from $\ell_1 > 0$ to $-f_1 < 0$, so player 1 prefers the opposite of t_2.
- (4) $\ell_1 < 0, f_1 < 0$: D_1 increases from $\ell_1 < 0$ to $-f_1 > 0$, so player 1 prefers to match t_2.

The same calculation for player 2, with $p = \Pr[t_1 = 1]$, gives $D_2(p) = (1-p)\ell_2 - pf_2$ and the symmetric four cases.

Pure equilibrium conditions. A pure profile is a Stage-1 NE iff each player's endpoint best-response condition holds. For example, $(1, 0)$ is a NE iff player 1 weakly prefers $t_1 = 1$ against $t_2 = 0$, i.e. $D_1(0) = \ell_1 \geq 0$, and player 2 weakly prefers $t_2 = 0$ against $t_1 = 1$, i.e. $D_2(1) = -f_2 \leq 0$, equivalently $f_2 \geq 0$. The other three pure profiles are identical endpoint checks, yielding the four displayed equivalences.

Fully mixed equilibria (general statement). A fully mixed equilibrium requires both players indifferent: $D_1(q) = D_2(p) = 0$ with $(p, q) \in (0, 1)^2$. The roots of the affine equations D_1 and D_2 are described by the indifference sets

$$Z_i = \begin{cases} \{\ell_i/(\ell_i + f_i)\}, & \ell_i + f_i \neq 0, \\ [0, 1], & \ell_i = f_i = 0, \\ \varnothing, & \ell_i + f_i = 0 \text{ and } \ell_i \neq 0, \end{cases}$$

so the fully mixed equilibrium set is the rectangle $(Z_2 \cap (0, 1)) \times (Z_1 \cap (0, 1))$.

In the nondegenerate case $\ell_i + f_i \neq 0$ for both players, both indifference sets are singletons, so the fully mixed equilibrium is unique whenever it exists; it is given by $p^* = \ell_2/(\ell_2 + f_2)$ and $q^* = \ell_1/(\ell_1 + f_1)$, valid iff both lie in $(0, 1)$. If either p^* or q^* falls outside $(0, 1)$, no fully mixed equilibrium exists and the equilibrium set reduces to the pure profiles already classified. In the fully degenerate case $\ell_1 = f_1 = \ell_2 = f_2 = 0$, every player is indifferent against every opponent mixture, so $(0, 1)^2$ is contained in the equilibrium set; this regime arises, e.g., in two-player constant-sum games (Corollary 1). Partial-degeneracy cases (one player nondegenerate, the other with $\ell_j = f_j = 0$) collapse to a one-parameter family of mixed equilibria along the degenerate player's indifference axis. □

B.2 Proof of Theorem 2 (Selection-Robust Equilibria)

We invoke (SA1, SA2). Recall the interval bounds

$$\underline{U}_i(t) = \inf_{x \in X(t)} x_i, \qquad \overline{U}_i(t) = \sup_{x \in X(t)} x_i.$$

Sufficiency. Suppose $\underline{U}_i(t) \geq \overline{U}_i(t'_i, t_{-i})$ for every player i and every unilateral deviation $t'_i \neq t_i$. Fix any admissible Stage-2 selection protocol χ (i.e., $\chi(s) \in \mathrm{Eq}_2(s)$ for all s and (SA2) holds).

Player i's selected payoff at t satisfies $U_i^\chi(t) \geq \underline{U}_i(t)$, while after the unilateral deviation (t_i', t_{-i}) the selected payoff satisfies $U_i^\chi(t_i', t_{-i}) \leq \overline{U}_i(t_i', t_{-i})$. Hence

$$U_i^\chi(t) \geq \underline{U}_i(t) \geq \overline{U}_i(t_i', t_{-i}) \geq U_i^\chi(t_i', t_{-i}),$$

so no player has a profitable Stage-1 deviation under any admissible χ, and t is selection-robust.

Necessity. Suppose the inequality fails for some player i and some deviation t_i', i.e. $\underline{U}_i(t) < \overline{U}_i(t_i', t_{-i})$. By definition of infimum and supremum, for any $\varepsilon > 0$ small enough there exist payoff vectors $x \in X(t)$ and $y \in X(t_i', t_{-i})$ such that

$$x_i < \underline{U}_i(t) + \varepsilon \quad \text{and} \quad y_i > \overline{U}_i(t_i', t_{-i}) - \varepsilon,$$

hence $x_i < y_i$ for ε small. Construct an admissible Stage-2 selection protocol χ as follows:

- Set $\chi(t)$ to a Stage-2 equilibrium with payoff vector x.
- Set $\chi(t_i', t_{-i})$ to a Stage-2 equilibrium with payoff vector y.
- For all other transparency profiles, choose any admissible Stage-2 equilibrium.
- If either $(0,0)$ or $(1,1)$ is among the profiles assigned, set the same payoff vector at both to enforce (SA2). The two relevant profiles for the deviation construction are t and (t_i', t_{-i}), and a unilateral transparency deviation never connects $(0,0)$ directly to $(1,1)$ (changing one coordinate of $(0,0)$ produces an asymmetric profile, not $(1,1)$), so the consistency constraint at $(0,0)$ and $(1,1)$ does not affect the assignments at t and (t_i', t_{-i}) except in the trivial case where $\{t, (t_i', t_{-i})\} \subseteq \{(0,0), (1,1)\}$, which is impossible for unilateral deviations.

Under this χ, player i's Stage-1 payoff at t is x_i and at the deviation is $y_i > x_i$, so the deviation is strictly profitable. Hence t is not selection-robust. □

B.3 Proof of Theorem 3 (Adaptive Wedge Certification)

We invoke (SA1, SA2, SA5, SA6).

Confidence sequence. Let $\widehat{U}_i^m(t)$ denote the empirical mean of player i's payoff after m samples of profile t, and let $S_i^m(t) = \sum_{j=1}^m (Y_i^j(t) - U_i^\chi(t))$ be the centered partial sum. By (SA5), $(S_i^m(t))_{m \geq 0}$ is a σ^2-sub-Gaussian martingale with respect to the natural filtration. For epoch $k \geq 0$, let $E_k = \{m : 2^k \leq m < 2^{k+1}\}$ and $h_k = \log(288(k+1)^2/\delta)$.

If $m \in E_k$ and $|\widehat{U}_i^m(t) - U_i^\chi(t)| > r(m)$, then since $r(m) = 2\sigma\sqrt{h_k/m}$ on E_k,

$$|S_i^m(t)| = m|\widehat{U}_i^m(t) - U_i^\chi(t)| > m \cdot 2\sigma\sqrt{h_k/m} = 2\sigma\sqrt{mh_k} \geq 2\sigma\sqrt{2^k h_k}.$$

The sub-Gaussian maximal inequality, obtained by applying Doob-Ville's martingale inequality to the exponential supermartingale $\exp(\lambda S_i^m(t) - \lambda^2\sigma^2 m/2)$ optimized over λ, gives

$$\Pr\left(\max_{s \leq 2^{k+1}} |S_i^s(t)| > 2\sigma\sqrt{2^k h_k}\right) \leq 2\exp(-h_k) = \frac{2\delta}{288(k+1)^2} = \frac{\delta}{144(k+1)^2}.$$

The algorithm samples three profiles (two players each), giving six sub-Gaussian sequences. A union bound over players, profiles, and epochs:

$$6\sum_{k \geq 0} \frac{\delta}{144(k+1)^2} = \frac{\delta}{24}\sum_{k \geq 0} \frac{1}{(k+1)^2} = \frac{\delta}{24} \cdot \frac{\pi^2}{6} < \delta.$$

Thus the event

$$\mathcal{E} := \{|\widehat{U}_i^m(t) - U_i^\chi(t)| \leq r(m) \text{ for all } i \in \{1,2\}, t \in \{(0,0),(1,0),(0,1)\}, m \geq 1\}$$

holds with probability at least $1 - \delta$.

Correctness on $\mathcal{E}$. Each wedge is a difference of two means, e.g. $\ell_1 = U_1^\chi(1,0) - U_1^\chi(0,0)$. On $\mathcal{E}$, the empirical wedge $\widehat{\ell}_1 = \widehat{U}_1(1,0) - \widehat{U}_1(0,0)$ satisfies $|\widehat{\ell}_1 - \ell_1| \leq r(N_{10}) + r(N_{00})$, so $\ell_1 \in I_{\ell_1}$. Hence whenever $0 \notin I_w$ for some wedge w, the empirical sign of w equals its true sign. By

Theorem 1, the four-wedge sign vector determines the strict pure-equilibrium classification; thus the classification returned at stopping is correct.

Sample bounds. Consider the two wedges ℓ_1, f_2 certified by profile $(1,0)$. Both have magnitude at least d_{10}. The wedge confidence radius for ℓ_1 is $r(N_{10}) + r(N_{00})$. Once

$$r(N_{10}) + r(N_{00}) < \frac{d_{10}}{2},$$

in particular $r(N_{10}) < d_{10}/2$ suffices on rounds where $r(N_{00})$ is also small, and the empirical magnitude of ℓ_1 is at least $d_{10} - r(N_{10}) - r(N_{00}) > d_{10}/2$, so the interval I_{ℓ_1} excludes zero. The same holds for f_2, so profile $(1,0)$ becomes inactive.

The dyadic radius $r(m) = 2\sigma\sqrt{h_k/m}$ for $m \in E_k$ satisfies $r(m) \le d/4$ once $m \ge 64\sigma^2 h_k/d^2$, which is implied by

$$m \ge C\frac{\sigma^2}{d^2}\left[\log\frac{1}{\delta} + \log\left(1 + \log\left(1 + \frac{\sigma^2}{d^2}\right)\right)\right]$$

for a universal constant C absorbing the dependence of h_k on $k = \lfloor \log_2 m \rfloor$. While profile $(1,0)$ is active, the algorithm samples both $(1,0)$ and $(0,0)$ in each loop iteration, so both radii reach the threshold simultaneously, giving the displayed bound on N_{10}. The argument for N_{01} is identical. The reference profile $(0,0)$ remains active until both asymmetric groups are inactive, so its sample count is controlled by the smallest bottleneck $d_0 = \min\{d_{10}, d_{01}\}$.

Lower bound. Fix any δ-correct adaptive algorithm in the Gaussian selected-payoff-table oracle model (where $Y_i(t) \sim \mathcal{N}(U_i^\theta(t), \sigma^2)$ and the algorithm adaptively samples any of the four profiles, with $(1,1)$ constrained to have the same payoff vector as $(0,0)$). For two payoff tables θ, θ', the standard adaptive transcript-KL identity gives

$$D_{\mathrm{KL}}(P_\theta \| P_{\theta'}) = \sum_t \mathbb{E}_\theta[N_t] \sum_{i=1}^{2} \frac{(U_i^\theta(t) - U_i^{\theta'}(t))^2}{2\sigma^2}.$$

If θ and θ' have different correct sign vectors, δ-correctness on both instances together with Le Cam's two-point inequality gives

$$D_{\mathrm{KL}}(P_\theta \| P_{\theta'}) \ge \mathrm{kl}(1-\delta, \delta) \ge \log\frac{1}{2.4\delta} = \Omega(\log(1/\delta))$$

for $\delta < 1/4$ (where (SA6) is invoked).

Asymmetric profile lower bounds. Let $a = (1,0)$. Choose the smaller of $|\ell_1|, |f_2|$; call its player index i and its signed value w, so $|w| = d_{10}$. Construct θ' by changing only $U_i^{\theta'}(a) = U_i^\theta(a) - (2+\eta)w$ for an arbitrarily small $\eta > 0$, chosen so that no untargeted wedge of θ' lands exactly at zero (the "bad" set of η-values is at most countable, so any sufficiently small generic η works). This preserves the simultaneous payoff (so (SA2) is preserved) and flips the chosen wedge from w to $-(1+\eta)w$, changing the sign vector. The KL identity gives

$$D_{\mathrm{KL}}(P_\theta \| P_{\theta'}) = \frac{\mathbb{E}_\theta[N_{10}]((2+\eta)d_{10})^2}{2\sigma^2} = \frac{(2+\eta)^2 d_{10}^2 \mathbb{E}_\theta[N_{10}]}{2\sigma^2} = \Theta\left(\frac{d_{10}^2 \mathbb{E}_\theta[N_{10}]}{\sigma^2}\right),$$

so $\mathbb{E}_\theta[N_{10}] = \Omega(\sigma^2 d_{10}^{-2}\log(1/\delta))$. The argument for N_{01} is identical with $a = (0,1)$.

Reference profile lower bound. Choose the globally smallest wedge over all four; write it as $w = U_i^\theta(a) - U_i^\theta(0,0)$ with $|w| = d_0$ and $a \in \{(1,0),(0,1)\}$. Construct θ' by changing both simultaneous-profile means $U_i^{\theta'}(0,0) = U_i^{\theta'}(1,1) = U_i^\theta(0,0) + (2+\eta)w$ for arbitrarily small generic $\eta > 0$, leaving asymmetric-profile means unchanged. This preserves (SA2), flips the chosen wedge from w to $-(1+\eta)w$, and (by genericity of η) keeps every other wedge nonzero. The KL identity gives

$$D_{\mathrm{KL}}(P_\theta \| P_{\theta'}) = \frac{\mathbb{E}_\theta[N_{\mathrm{sim}}]((2+\eta)d_0)^2}{2\sigma^2} = \Theta\left(\frac{d_0^2 \mathbb{E}_\theta[N_{\mathrm{sim}}]}{\sigma^2}\right),$$

so $\mathbb{E}_\theta[N_{\mathrm{sim}}] = \Omega(\sigma^2 d_0^{-2}\log(1/\delta))$ where $N_{\mathrm{sim}} = N_{00} + N_{11}$.

Summing the three necessary profile-specific lower bounds proves the certificate-complexity claim

$$\Theta(\sigma^2(d_0^{-2} + d_{10}^{-2} + d_{01}^{-2})\log(1/\delta))$$

up to the anytime logarithmic factor incurred by the dyadic stitching. □

C Deferred from Main Text

C.1 Special Game Classes

Definition 8 (Constant-Sum Game)**.** *G is constant-sum if $\sum_{i \in N} u_i(s) = c$ for all $s \in S$, for some constant c. Zero-sum games are the special case $c = 0$.*

Definition 9 (Pure Coordination Game)**.** *G is a pure coordination game if there exists $v : S \to \mathbb{R}$ such that $u_i(s) = v(s)$ for all $i \in N$ and all $s \in S$.*

Definition 10 (Mixed-Motive Game)**.** *G is mixed-motive if it is neither constant-sum nor pure coordination.*

C.2 Benchmark Class Corollaries

The two benchmark classes calibrate the wedge theorem.

Corollary 1 (Constant-sum games)**.** *In a two-player constant-sum game, the minimax theorem implies $L_i = F_i = N_i$ for both players under the baseline PTG timing model. Hence all wedges vanish, every transparency profile is payoff-equivalent, and every Stage-1 profile is an equilibrium. If a positive transparency cost is added, mutual opacity is the unique Stage-1 equilibrium.*

Corollary 2 (Pure coordination games)**.** *In a two-player identical-interest game with common payoff v, unilateral transparency has an SPE achieving the global maximum $v^* = \max_s v(s)$. If simultaneous profiles select a payoff-dominant Nash equilibrium, then $L_i = F_i = N_i = v^*$ for both players, all wedges vanish, and transparency is payoff-irrelevant. Without payoff-dominant selection, the conclusion is selection-dependent and is evaluated using Theorem 2.*

C.3 SSE Leader-Wedge Nonnegativity

Lemma 1 (SSE leader-wedge nonnegativity)**.** *Let χ be the default optimistic SSE protocol at asymmetric transparency profiles, paired with any simultaneous-Nash selection at the simultaneous profiles. Then for every two-player base game G and both players $i \in \{1, 2\}$,*

$$\ell_i = L_i - N_i \geq 0.$$

Proof. Let $\sigma^* = (\sigma_i^*, \sigma_{-i}^*)$ be the simultaneous Nash equilibrium that χ selects at $(0, 0)$ (and hence, by (SA2), at $(1, 1)$), so $N_i = u_i(\sigma^*)$. Suppose player i is the sole transparent leader and consider the candidate commitment σ_i^*. Because σ_{-i}^* is a best response to σ_i^* (Nash condition), σ_{-i}^* lies in the follower's best-response set $BR_{-i}(\sigma_i^*)$. Optimistic SSE breaks follower ties in favor of the leader, so the follower response selected at the commitment σ_i^* delivers payoff to player i at least $u_i(\sigma_i^*, \sigma_{-i}^*) = N_i$. Since L_i is the maximum of the leader payoff over all commitments, $L_i \geq N_i$, i.e. $\ell_i \geq 0$. □

Corollary 3 (Sign regimes feasible under default SSE)**.** *Under the default optimistic SSE protocol with no transparency costs, the wedge-sign regimes from Theorem 1 satisfy $\ell_i \geq 0$ for both players. Consequently:*

1. Strict opacity dominance *($\ell_i < 0$ and $f_i > 0$) and* matching *($\ell_i < 0$ and $f_i < 0$) regimes from cases (2) and (4) of Theorem 1 are not attainable.*

2. *Mutual opacity $(0, 0)$ can be a Stage-1 equilibrium* only weakly*, when $\ell_1 = \ell_2 = 0$. Strict mutual-opacity dominance does not occur.*

3. *Achieving $\ell_i < 0$ requires leaving the default optimistic-SSE asymmetric-profile protocol (e.g., pessimistic or non-optimistic follower tie-breaking), changing the simultaneous benchmark away from any Nash-consistent simultaneous selection (e.g., a non-equilibrium simultaneous reference payoff), or adding direct transparency costs/incentives (the augmented PTG of Appendix C.10.2). Merely selecting a* different *simultaneous Nash equilibrium does* ***not*** *make $\ell_i < 0$ under optimistic SSE: the same proof of Lemma 1 applied with the chosen σ^* shows $L_i \geq u_i(\sigma^*) = N_i$ for every Nash σ^*.*

The four-case algebraic statement in Theorem 1 remains valid for the broader class of simultaneous-profile-consistent selection protocols allowed by (SA2); Lemma 1 only restricts which sign vectors a specific *protocol can realize.*

C.4 Strict Mutual-Transparency Example

Example 1 (Strict mutual-transparency equilibrium)**.** *The following symmetric 2×2 base game gives a strict separation in which mutual transparency $(1,1)$ is the **unique** Stage-1 Nash equilibrium:*

$$U_1 = \begin{pmatrix} 4 & 1 \\ 5 & 0 \end{pmatrix}, \quad U_2 = \begin{pmatrix} 4 & 5 \\ 1 & 0 \end{pmatrix}.$$

*The simultaneous game has three Nash equilibria: two pure profiles $(\textit{Row 1}, \textit{Column 2})$ and $(\textit{Row 2}, \textit{Column 1})$, and a unique fully mixed equilibrium $p^{NE} = q^{NE} = 1/2$ with payoffs $N_1 = N_2 = 5/2$. We fix the simultaneous selection of χ at $(0,0)$ and $(1,1)$ to be this **mixed** Nash equilibrium; pure-equilibrium simultaneous selections produce different wedges and are not discussed here.*

Profile $(1,0)$: *Player 1 as SSE leader commits to Row 2, which makes Column 1 the unique follower best response, yielding $(L_1, F_2) = (5,1)$.* Profile $(0,1)$: *Player 2 as SSE leader commits to Column 2, which makes Row 1 the unique follower best response, yielding $(F_1, L_2) = (1,5)$.*

Hence the wedges are

$$\ell_1 = \ell_2 = 5 - \tfrac{5}{2} = \tfrac{5}{2} > 0, \qquad f_1 = f_2 = 1 - \tfrac{5}{2} = -\tfrac{3}{2} < 0,$$

so each player satisfies the strict transparency-dominance regime $\ell_i > 0, f_i < 0$ of Theorem 1 (case 1). The induced Stage-1 payoff matrices are

$$\Gamma_1 = \begin{pmatrix} 5/2 & 1 \\ 5 & 5/2 \end{pmatrix}, \quad \Gamma_2 = \begin{pmatrix} 5/2 & 5 \\ 1 & 5/2 \end{pmatrix},$$

and direct inspection (or Theorem 1 case 1) confirms that $(1,1)$ is the unique Stage-1 Nash equilibrium. This is a clean Cournot-style instance of the wedge mechanism, consistent with the SSE leader-wedge nonnegativity of Lemma 1 ($\ell_1 = \ell_2 = 5/2 \geq 0$).

C.5 Compact Wedge Illustration

The wedge theorem replaces long case analyses with sign checks. In Battle of the Sexes with mixed simultaneous-equilibrium selection, the leader and follower payoffs in the two asymmetric transparency profiles both exceed the simultaneous mixed payoff. Thus both leader and follower wedges are positive, and Theorem 1 gives asymmetric transparency equilibria $(1,0)$ and $(0,1)$. The example is useful because it shows that mutual transparency and mutual opacity can have the same selected payoff while both fail Stage-1 stability for the same reason: each player wants the opposite transparency status from the opponent.

C.6 Welfare Notation

For later examples we use social welfare

$$W^{\chi}(t) := \sum_i U_i^{\chi}(t).$$

When comparing transparency policies, the additive welfare gap

$$\mathrm{AOG}_{\chi}(G) := \max_{t \in \{0,1\}^n} W^{\chi}(t) - \min_{\tau^* \in \mathrm{Eq}_{\chi}(\mathrm{PTG})} \mathbb{E}_{t \sim \tau^*}[W^{\chi}(t)]$$

is invariant to common payoff translations. A ratio version, the Price of Opacity, can also be reported under an explicit payoff normalization; the results below use the additive gap.

C.7 No-Regret Transparency Learning

Consider repeated play for T rounds of a fixed two-player PTG under selection protocol χ. In round r, each player chooses $t_i^r \in \{0,1\}$, receives payoff $U_i^{\chi}(t^r)$, and observes the full-information transparency feedback above. Player i has external regret $R_i(T)$ if

$$\max_{a_i \in \{0,1\}} \sum_{r=1}^{T} U_i^{\chi}(a_i, t_{-i}^r) - \sum_{r=1}^{T} U_i^{\chi}(t_i^r, t_{-i}^r) \leq R_i(T).$$

Let $\widehat{\mu}_T$ be the empirical distribution of the transparency profiles $(t^1, \ldots, t^T)$.

Theorem 4 (No-regret transparency learning)**.** *Assume (SA1, SA2, SA4): the selected payoffs lie in* $[0,1]$ *and the learner observes the full-information transparency feedback. If both players have external regret* $R_i(T)$*, then* $\widehat{\mu}_T$ *is an* ε_T*-coarse correlated equilibrium of the induced transparency game with player-wise errors* $\varepsilon_i = R_i(T)/T$*:*

$$\mathbb{E}_{t\sim\widehat{\mu}_T}[U_i^{\chi}(t)] \geq \max_{a_i\in\{0,1\}} \mathbb{E}_{t\sim\widehat{\mu}_T}[U_i^{\chi}(a_i, t_{-i})] - \frac{R_i(T)}{T}.$$

Moreover, suppose there is a strict dominant transparency action t_i^* *for each player with margin* $\gamma_i > 0$*:*

$$U_i^{\chi}(t_i^*, t_{-i}) \geq U_i^{\chi}(1 - t_i^*, t_{-i}) + \gamma_i \qquad \forall t_{-i} \in \{0,1\}.$$

Then the number $M_i(T)$ *of rounds in which player* i *does not play* t_i^* *satisfies*

$$M_i(T) \leq \frac{R_i(T)}{\gamma_i},$$

and the fraction of rounds outside the dominant transparency profile t^* *is at most*

$$\frac{1}{T}\left(\frac{R_1(T)}{\gamma_1} + \frac{R_2(T)}{\gamma_2}\right).$$

In particular, under the full-information feedback model of (SA4), Hedge over the two transparency actions yields $R_i(T) \leq \sqrt{(T/2)\log 2}$*, so strict wedge-dominance regimes converge to the dominant transparency profile at rate* $O(T^{-1/2}/\gamma)$*. Under bandit feedback (only* $U_i^{\chi}(t^r)$ *observed), EXP3 achieves* $\mathbb{E}[R_i(T)] = O(\sqrt{T\log 2})$*, and the CCE and mistake-fraction conclusions hold in expectation with the same* $O(T^{-1/2}/\gamma)$ *rate up to constants.*

Proof. We invoke (SA1, SA2, SA4).

CCE bound. By the external-regret hypothesis, for every $a_i \in \{0,1\}$ and every player i,

$$\sum_{r=1}^{T} U_i^{\chi}(a_i, t_{-i}^r) - \sum_{r=1}^{T} U_i^{\chi}(t_i^r, t_{-i}^r) \leq R_i(T).$$

Dividing by T and rewriting each time average as the expectation under the empirical distribution $\widehat{\mu}_T$ of the realized profile sequence $(t^r)_{r=1}^T$:

$$\mathbb{E}_{t\sim\widehat{\mu}_T}[U_i^{\chi}(a_i, t_{-i})] - \mathbb{E}_{t\sim\widehat{\mu}_T}[U_i^{\chi}(t)] \leq \frac{R_i(T)}{T}.$$

Rearranging gives the displayed ε_i-CCE inequality with $\varepsilon_i = R_i(T)/T$.

Mistake bound. Fix player i and the strict dominant transparency action t_i^* with margin $\gamma_i > 0$. Compare player i's realized cumulative payoff to the comparator cumulative payoff from playing t_i^* on every round, while the opponents' actions t_{-i}^r are fixed. On each round r where player i plays the dominant action ($t_i^r = t_i^*$), the per-round difference is zero. On each round where she plays $1 - t_i^*$, the strict-margin assumption gives

$$U_i^{\chi}(t_i^*, t_{-i}^r) - U_i^{\chi}(1 - t_i^*, t_{-i}^r) \geq \gamma_i.$$

Summing over rounds,

$$\sum_{r=1}^{T} U_i^{\chi}(t_i^*, t_{-i}^r) - \sum_{r=1}^{T} U_i^{\chi}(t_i^r, t_{-i}^r) \geq \gamma_i M_i(T).$$

The left-hand side is at most $R_i(T)$ by the external-regret guarantee applied to the comparator action t_i^*. Hence $\gamma_i M_i(T) \leq R_i(T)$, i.e. $M_i(T) \leq R_i(T)/\gamma_i$.

Profile bound. A round is outside the dominant transparency profile $t^* = (t_1^*, t_2^*)$ iff at least one player makes a mistake. By the union bound,

$$\#\{r : t^r \neq t^*\} \leq M_1(T) + M_2(T) \leq \frac{R_1(T)}{\gamma_1} + \frac{R_2(T)}{\gamma_2}.$$

Dividing by T gives the stated profile-fraction bound.

Hedge rate. Under (SA4), payoffs lie in $[0,1]$ and player i observes both counterfactual payoffs $U_i^X(0,t_{-i}^r), U_i^X(1,t_{-i}^r)$ at the end of round r. This is exactly the full-information experts setting on $K=2$ actions: the per-round loss vector $(1-U_i^X(0,t_{-i}^r), 1-U_i^X(1,t_{-i}^r))$ is fully observed. The Hedge algorithm with learning rate $\eta=\sqrt{(8\log K)/T}$ achieves the standard regret bound $R_i(T)\le\sqrt{(T/2)\log K}$; with $K=2$, $R_i(T)\le\sqrt{(T/2)\log 2}$. Substituting into the mistake-fraction bound yields the $O(T^{-1/2}/\gamma)$ convergence rate.

Under bandit feedback (only $U_i^X(t^r)$ observed), EXP3 over two actions achieves the in-expectation regret bound $\mathbb{E}[R_i(T)]\le 2\sqrt{T\log 2}$; the CCE inequality then holds in expectation, the mistake-bound argument applies in expectation by linearity, and the mistake-fraction rate becomes $O(T^{-1/2}/\gamma)$ with worse constants. □

In wedge notation, the strict dominant-transparency case is $\ell_i>0$ and $f_i<0$ for both players, with $\gamma_i=\min\{\ell_i,-f_i\}$. Thus in the Cournot-style regime, no-regret learners reach $(1,1)$ with a finite mistake bound controlled by the wedge margin.

C.8 Targeted Equilibrium Certification

Full wedge-sign recovery is sometimes more information than the learner needs. If a regulator or platform wants to certify a proposed strict transparency equilibrium t, only the two unilateral-deviation inequalities at t matter. The wedge representation turns this into a smaller comparison graph.

For a pure transparency profile $t\in\{0,1\}^2$, let $d_i(t)=(1-t_i,t_{-i})$ be player i's unilateral deviation. Let $\rho(1,1)=(0,0)$ and $\rho(s)=s$ for all other profiles, using simultaneous-profile consistency to represent $(1,1)$ by $(0,0)$. Define the strict equilibrium gaps

$$g_i(t):=U_i^X(\rho(t))-U_i^X(\rho(d_i(t))),\qquad i=1,2.$$

For each representative profile $s\in C(t):=\{\rho(t),\rho(d_1(t)),\rho(d_2(t))\}$, define its incident bottleneck

$$b_s(t):=\min\{g_i(t):s\in\{\rho(t),\rho(d_i(t))\}\}.$$

When t is a strict equilibrium, all $g_i(t)$ and all $b_s(t)$ are positive.

Theorem 5 (Local certification of a proposed transparency equilibrium)**.** *Fix $\delta\in(0,1/4)$, a two-player PTG under a simultaneous-profile-consistent selection protocol, and a proposed pure profile t with strict gaps $g_1(t),g_2(t)>0$. There is an adaptive LOCAL-CERTIFY algorithm that samples only profiles in $C(t)$ and certifies that t is a strict Stage-1 equilibrium with probability at least $1-\delta$. On this event, for every $s\in C(t)$,*

$$N_s=O\left(\frac{\sigma^2}{b_s(t)^2}\left[\log\frac{1}{\delta}+\log\left(1+\log\left(1+\frac{\sigma^2}{b_s(t)^2}\right)\right)\right]\right),$$

where $N_{(0,0)}$ denotes pulls of the representative simultaneous profile used by the algorithm.

Call a local certifier δ-correct if, on every instance where t is a strict equilibrium, it certifies with probability at least $1-\delta$, and on every instance where at least one tested deviation gap is nonpositive, it certifies with probability at most δ.

Conversely, in the Gaussian selected-payoff-table oracle model of Theorem 3 (which is a lower bound over the payoff-table class, not a realization lower bound over finite base games), any δ-correct local certifier for t must satisfy the profile-level lower bound

$$\mathbb{E}[\bar N_s]=\Omega\left(\frac{\sigma^2}{b_s(t)^2}\log\frac{1}{\delta}\right)\qquad\text{for every } s\in C(t),$$

where $\bar N_s=N_s$ for asymmetric profiles and $\bar N_{(0,0)}=N_{00}+N_{11}$ for the simultaneous representative.

Proof. We invoke (SA1, SA2, SA5, SA6).

Algorithm. LOCAL-CERTIFY is WEDGE-CERTIFY restricted to the profile graph $C(t)=\{\rho(t),\rho(d_1(t)),\rho(d_2(t))\}$ and to the two unilateral-deviation gaps $g_1(t),g_2(t)$. It pulls every profile

in $C(t)$ once, maintains dyadic confidence intervals for the two gaps using the same radius $r(\cdot)$, and stops when both lower confidence bounds are positive.

Confidence event and correctness. Restrict the event $\mathcal{E}$ from the proof of Theorem 3 to the profiles in $C(t)$; the union bound now sums over at most three profiles times two players, giving the same $1-\delta$ probability. On this event, each gap $g_i(t)$ is contained in its empirical confidence interval. If both lower confidence bounds for $g_1(t)$ and $g_2(t)$ are positive, then both true gaps are positive, so both unilateral deviations from t are strictly unprofitable, certifying that t is a strict Stage-1 NE.

Sample bounds. A profile $s \in C(t)$ occurs only in the gap comparisons whose deviation edge is incident to s. The incident bottleneck $b_s(t)$ is the smallest such gap. The same radius inversion as in the proof of Theorem 3 shows that once $r(N_s) \le b_s(t)/4$ holds for every active incident profile s, every incident gap has empirical magnitude exceeding $b_s(t)/2$ and confidence radius below $b_s(t)/2$, so the gap is certified. The threshold

$$N_s \ge C\frac{\sigma^2}{b_s(t)^2}\left[\log\frac{1}{\delta} + \log\left(1+\log\left(1+\frac{\sigma^2}{b_s(t)^2}\right)\right)\right]$$

gives the displayed upper bound.

Lower bound. Fix any $s \in C(t)$ and apply the Gaussian change-of-measure argument of the proof of Theorem 3 to one incident gap attaining $b_s(t)$, say $g_i(t)$. Use a perturbation magnitude $(2+\eta)g_i(t)$ for arbitrarily small generic $\eta > 0$ chosen so that no other gap of θ' becomes exactly zero (a countable bad set).

- If $s = \rho(d_i(t))$, perturb only player i's payoff at the deviation representative by $+(2+\eta)g_i(t)$.
- If $s = \rho(t)$, perturb player i's payoff at the equilibrium representative by $-(2+\eta)g_i(t)$.
- If $s = (0,0)$ (i.e., $\rho(t) = (0,0)$ and $t = (1,1)$ for instance), apply the same change to both simultaneous profiles $(0,0)$ and $(1,1)$ to preserve (SA2).

In every case, the chosen gap $g_i(t)$ flips sign (from $+g_i(t)$ to $-(1+\eta)g_i(t)$), so a δ-correct local certifier must distinguish the two instances. The transcript-KL identity gives

$$D_{\mathrm{KL}}(P_\theta \| P_{\theta'}) = \frac{((2+\eta)b_s(t))^2\mathbb{E}[\bar{N}_s]}{2\sigma^2} = \Theta\left(\frac{b_s(t)^2\mathbb{E}[\bar{N}_s]}{\sigma^2}\right)$$

where $\bar{N}_s = N_s$ for asymmetric s and $\bar{N}_{(0,0)} = N_{00} + N_{11}$. Combined with $\mathrm{kl}(1-\delta,\delta) = \Omega(\log(1/\delta))$ from (SA6), this yields the displayed per-profile lower bound. □

C.9 Battle of the Sexes Certification Example

Example 2 (Certification budgets in Battle of the Sexes)**.** *In the standard Battle of the Sexes game with mixed simultaneous-equilibrium selection, each player's simultaneous payoff is $N_i = 2/3$. A sole leader obtains payoff 2, while a sole follower obtains payoff 1. Hence*

$$\ell_i = 2 - \frac{2}{3} = \frac{4}{3}, \qquad f_i = 1 - \frac{2}{3} = \frac{1}{3}.$$

Full wedge certification has $d_{10} = d_{01} = d_0 = 1/3$, so Theorem 3 samples all three informative profiles at the bottleneck scale $9\sigma^2$ up to logarithmic factors. To certify the proposed asymmetric equilibrium $(1,0)$, Theorem 5 only tests $\ell_1 > 0$ and $f_2 > 0$. The local certificate graph is $\{(0,0),(1,0)\}$ with bottleneck $1/3$ on both profiles, so the learner avoids sampling $(0,1)$ entirely while retaining profile-level matching upper and lower bounds.

C.10 Computational Complexity

We establish the computational landscape for PTG equilibrium computation.

C.10.1 Algorithm for 2×2 PTG Equilibrium

Algorithm 2 PTG-Equilibrium-2x2

1: **Input:** Payoff matrices $U_1, U_2 \in \mathbb{R}^{2\times 2}$
2: **Output:** Equilibrium Stage 1 strategy profile $\tau^* = (p, q)$ where $p = \Pr[t_1 = 1]$ and $q = \Pr[t_2 = 1]$, plus the four Stage 2 payoff vectors $\pi^{(0,0)}, \pi^{(1,0)}, \pi^{(0,1)}, \pi^{(1,1)}$

Step 1: Compute Stage-2 Payoffs for All Transparency Profiles

1.1. Profile $(0, 0)$: enumerate all supports of the 2×2 simultaneous game, solve the associated indifference equations, discard candidates violating best-response inequalities, and apply the fixed protocol χ to the resulting Nash equilibrium set. Let $\pi^{(0,0)}$ be the selected payoff vector.

1.2. Profile $(1, 0)$: compute the SSE leader value and the **full set of leader-optimal SSE outcomes** for player 1 by the **standard one-LP-per-follower-pure-action** formulation. For each follower pure action $j \in \{1, 2\}$, impose the linear inequalities on player 1's mixed commitment $p \in [0, 1]$ that make j a follower best response, obtaining a closed interval $I_j \subseteq [0, 1]$ (empty if no commitment makes j a best response). The leader payoff conditional on follower action j is affine in p, so its maximum over I_j is attained at an endpoint of I_j.

(a) **Leader value.** Apply optimistic SSE tie-breaking at any shared endpoint (where the follower is indifferent and SSE selects the follower action that maximizes player 1's payoff). Let

$$L_1 := \max_{j\in\{1,2\}} \max_{p\in I_j} u_1(p, j)$$

be the SSE leader value.

(b) **Leader-optimal SSE outcome set.** Under the nondegeneracy hypothesis of Theorem 6, collect the finite endpoint set $\mathcal{O}^{(1,0)} \subseteq [0, 1] \times \{1, 2\}$ of all $(p^\star, j^\star)$ that attain $u_1(p^\star, j^\star) = L_1$ after optimistic follower tie-breaking, where $p^\star$ ranges over the at-most-four endpoints of $I_1 \cup I_2$. Each element of $\mathcal{O}^{(1,0)}$ is a leader-optimal optimistic-SSE outcome; different elements can give different follower payoffs $u_2(p^\star, j^\star)$. (Without nondegeneracy, for example when player 1's leader payoff is flat across an interval that the follower best-responds to with a single pure action, the leader-optimal SSE outcome set can be a continuum, and χ must instead operate on a polyhedral interval description, as in Remark 2; the same applies to Step 1.3.)

(c) **Apply χ.** Invoke the fixed protocol χ on $\mathcal{O}^{(1,0)}$ to select a single outcome $(p^\star, j^\star)$, and set $\pi^{(1,0)} = (L_1, F_2)$ where $F_2 := u_2(p^\star, j^\star)$. (Per Section 2.2, χ includes a deterministic tie-breaking rule over $\mathcal{O}^{(1,0)}$.)

1.3. Profile $(0, 1)$: Symmetric to Step 1.2 with Player 2 as leader. Compute the analogous outcome set $\mathcal{O}^{(0,1)}$ and let $\pi^{(0,1)} = (F_1, L_2)$ be the payoff vector selected by χ on $\mathcal{O}^{(0,1)}$.

1.4. Profile $(1, 1)$: By Remark 1, $\pi^{(1,1)} := \pi^{(0,0)}$ under the same selection rule.

Step 2: Construct Stage 1 Game

Form 2×2 game Γ with payoff matrices:

$$\Gamma_1 = \begin{pmatrix} \pi_1^{(0,0)} & \pi_1^{(0,1)} \\ \pi_1^{(1,0)} & \pi_1^{(1,1)} \end{pmatrix}, \quad \Gamma_2 = \begin{pmatrix} \pi_2^{(0,0)} & \pi_2^{(0,1)} \\ \pi_2^{(1,0)} & \pi_2^{(1,1)} \end{pmatrix}$$

Step 3: Solve Stage-1 Game

3.1. Check the four pure transparency profiles by verifying the best-response inequalities.

3.2. If a pure equilibrium is found, return one such equilibrium. If no pure equilibrium exists, compute the mixed equilibrium by solving the two affine indifference equations from Theorem 1:

$$p = \frac{\ell_2}{\ell_2 + f_2}, \qquad q = \frac{\ell_1}{\ell_1 + f_1},$$

where $\ell_i = L_i - \pi_i^{(0,0)}$ and $f_i = F_i - \pi_i^{(0,0)}$. If a denominator is zero or a candidate lies outside $[0, 1]$, solve the remaining boundary supports by direct support enumeration of the 2×2 Stage-1 game.

Step 4: Return Result

Return equilibrium Stage-1 strategy $\tau^* = (p, q)$ and the four selected Stage-2 payoff vectors.

Theorem 6 (Correctness and Complexity)**.** *Assume (SA1, SA2, SA7) and the following two additional hypotheses:*

1. (Nondegeneracy.) *The simultaneous 2×2 base game G is **nondegenerate** in the standard bimatrix-game sense; equivalently, no mixed strategy of one player makes the other player indifferent over a strictly larger support than that mixed strategy itself, so the Nash equilibrium set of the simultaneous base game is finite.*
2. (Tie-breaking computability.) *The protocol χ is computable in polynomial time **both** on the (constant-size) set of simultaneous Nash equilibria returned by support enumeration and on the (constant-size) leader-optimal SSE outcome sets $\mathcal{O}^{(1,0)}, \mathcal{O}^{(0,1)}$ produced in Steps 1.2-1.3.*

Then Algorithm 2 correctly computes a PTG equilibrium under χ, and its bit complexity is polynomial in the input bit-length.

Remark 2 (Degenerate base games)**.** *Without the nondegeneracy assumption, the simultaneous Nash equilibrium set can be a continuum (e.g., the all-zero 2×2 game has equilibrium set $[0, 1]^2$). Algorithm 2 still terminates if χ is upgraded to operate on a polyhedral / semialgebraic equilibrium description and to return one selected payoff vector in polynomial time on such a description. We adopt the simpler nondegeneracy assumption in the main text and treat the polyhedral-selection variant as a straightforward extension.*

Proof. We invoke (SA1, SA2, SA7), the nondegeneracy assumption of Theorem 6, and the tie-breaking-computability assumption that χ is computable in polynomial time on the simultaneous Nash set returned by support enumeration **and** on the leader-optimal SSE outcome sets $\mathcal{O}^{(1,0)}, \mathcal{O}^{(0,1)}$.

Step 1.1 correctness. Every Nash equilibrium of a 2×2 bimatrix game has supports drawn from one of $3 \times 3 = 9$ support pairs. For each support pair, the indifference equations on the chosen supports, together with the best-response inequalities on the unchosen actions, form a linear system with rational coefficients (under (SA7)). Under the nondegeneracy assumption of Theorem 6, this enumeration returns a *finite* Nash equilibrium set; standard exact support enumeration solves each system in constant time, and the protocol χ is then invoked once on this constant-size set to select $\pi^{(0,0)}$. (Without nondegeneracy, e.g. for the all-zero 2×2 game, the equilibrium set can be a continuum and χ must instead operate on a polyhedral description, as noted in Remark 2.)

Step 1.2 correctness (one LP per follower pure action; χ resolves leader-optimal ties). Use the standard Stackelberg-LP formulation specialized to two follower actions. For each follower pure action $j \in \{1, 2\}$, the linear inequality

$$\Phi_j(p) := p(b_{1j} - b_{1,3-j}) + (1 - p)(b_{2j} - b_{2,3-j}) \geq 0$$

characterizes the leader commitments $p \in [0, 1]$ for which j is a follower best response. The set $I_j = \{p \in [0, 1] : \Phi_j(p) \geq 0\}$ is a closed sub-interval of $[0, 1]$ (empty if no commitment makes j a best response), since Φ_j is affine in p. The leader payoff conditional on follower action j is

$$u_1(p, j) = p \cdot a_{1j} + (1 - p) \cdot a_{2j},$$

again affine in p, so its maximum over the closed interval I_j is attained at an endpoint of I_j. The SSE commitment value for player 1 is

$$L_1 = \max_{j \in \{1,2\}} \max_{p \in I_j} u_1(p, j),$$

where ties at any shared endpoint of I_1 and I_2 are resolved by optimistic SSE in the leader's favor. Each inner maximum is the larger of the two endpoint evaluations, and the outer maximum is then computed by direct comparison.

Optimistic SSE pins the leader value L_1 uniquely, but several leader-optimal outcome pairs $(p^\star, j^\star)$ can attain $u_1(p^\star, j^\star) = L_1$ (for example, two distinct endpoints in $I_1 \cup I_2$), and these can induce *different* follower payoffs $u_2(p^\star, j^\star)$. (A simple instance: $U_1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), U_2 = \left(\begin{smallmatrix} 0 & 1 \\ 4 & 3 \end{smallmatrix}\right)$ gives leader-optimal SSE outcomes $(p, j) = (0, 1)$ with follower payoff 4 and $(p, j) = (1, 2)$ with follower payoff 1, both at $L_1 = 1$.) Algorithm 2 therefore collects the finite leader-optimal SSE outcome set

$$\mathcal{O}^{(1,0)} = \{(p^\star, j^\star) \in \{\text{endpoints of } I_1 \cup I_2\} \times \{1, 2\} : u_1(p^\star, j^\star) = L_1 \text{ after optimistic follower tie-breaking}\}$$

and invokes χ on $\mathcal{O}^{(1,0)}$ to select a single $(p^\star, j^\star)$, setting $F_2 := u_2(p^\star, j^\star)$. Since $\mathcal{O}^{(1,0)}$ has size at most a constant (each I_j has at most two endpoints), χ is invoked on a constant-size finite set, and (SA7) gives polynomial bit complexity for the invocation. Hence $\pi^{(1,0)} = (L_1, F_2)$ is computed exactly. (In the degenerate Stackelberg case where $\Phi_1 \equiv 0$, both follower actions are best responses on all of $[0, 1]$ and SSE selects the action giving the larger leader payoff at each p, which is again handled by endpoint comparison and ties absorbed into $\mathcal{O}^{(1,0)}$.)

Step 1.3 and Step 1.4. Step 1.3 is identical to Step 1.2 with the players' roles swapped. Step 1.4 invokes Remark 1 directly: $\pi^{(1,1)} = \pi^{(0,0)}$ under the consistency convention (SA2) of χ.

Step 3 correctness. After Step 1, the PTG is reduced to the 2×2 Stage-1 normal-form game $\Gamma = (\Gamma_1, \Gamma_2)$. Step 3.1 checks the four pure transparency profiles by verifying the best-response inequalities; if any is satisfied, return that pure equilibrium. If no pure equilibrium exists, then by Theorem 1 the unique mixed equilibrium is interior and given by

$$p = \frac{\ell_2}{\ell_2 + f_2}, \qquad q = \frac{\ell_1}{\ell_1 + f_1},$$

which Step 3.2 computes directly. Degenerate or boundary cases (zero denominators, probabilities outside $[0, 1]$) are handled by support enumeration of the 2×2 Stage-1 game, which is again exact. Thus the returned τ^* is a Nash equilibrium of Γ and therefore a PTG equilibrium under χ.

Complexity. Each of Steps 1.1-1.4 performs a constant number of rational additions, multiplications, comparisons, and one-variable linear solves. Numerators and denominators produced by these operations have bit-length polynomial in the input bit-length (the bit-length of the rational entries of U_1, U_2). The selection protocol χ is invoked only on constant-size equilibrium sets; under (SA7), each invocation runs in time polynomial in the input bit-length. Step 3 likewise performs a constant number of rational arithmetic operations. Hence Algorithm 2 has polynomial bit complexity. □

C.10.2 Inherited Hardness and Pessimistic Selection

The positive computational result is the exact 2×2 algorithm above. We also record how augmented variants inherit known hardness from embedded Stage-2 subproblems.

The hardness statements below concern the **full PTG equilibrium-output problem**: the algorithm must output a Stage-1 transparency equilibrium *together with* the selected Stage-2 equilibrium (or its payoff/strategy) at every transparency profile in the Stage-1 equilibrium support, or at minimum at the realized profile that the Stage-1 equilibrium plays with positive probability. This is the natural decision problem for an analyst or simulator that needs to predict Stage-2 play, not only the transparency choice.

First, under the full-output requirement, augmented PTGs with exogenous transparency costs inherit the hardness of the base game's simultaneous Nash problem: a sufficiently large opacity-favoring cost makes $(0, 0)$ the unique Stage-1 profile, and outputting the full PTG equilibrium then requires computing a Nash equilibrium of the base bimatrix game at $(0, 0)$. This embedding gives PPAD-hardness for the full-output problem (Chen et al., 2009; Daskalakis et al., 2009). If the computational task is only to output the Stage-1 transparency action (ignoring the Stage-2 equilibrium that will subsequently be played), the same forced-opacity reduction is trivial (the answer is $(0, 0)$), and no PPAD-hardness follows from this construction.

Second, augmented multi-follower PTGs inherit the known NP-hardness of pessimistic multi-follower Stackelberg equilibrium computation: large exogenous incentives can force the transparency profile with one leader and multiple followers, leaving the fixed-profile Stackelberg subproblem unchanged (Coniglio et al., 2017). Again, the hardness is inherited from the selected Stage-2 subproblem.

Finally, pessimistic multi-follower selection can destroy totality: the leader's pessimistic objective can have a supremum that is not attained, a known phenomenon in pessimistic Stackelberg models.

We therefore treat pessimistic multi-follower PTGs as requiring either an ε-equilibrium formulation or an explicit existence assumption.

Table 1: Computational Boundary

Setting	Status
Two-player 2×2 PTG (Stage-1 + Stage-2 output)	Polynomial time by Algorithm 2
Augmented two-player PTG with costs, **full equilibrium-output problem** (Stage-1 + Stage-2)	Inherits PPAD-hardness from forced simultaneous Nash
Augmented two-player PTG with costs, **Stage-1-only output under forced opacity**	Trivial in that reduction (output is $(0, 0)$); no hardness follows from this construction
Augmented multi-follower PTG, full equilibrium-output problem	Inherits NP-hardness from fixed-profile pessimistic Stackelberg
Pure-pessimistic multi-follower selection	May be non-total without ε-relaxation

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction clearly state the main claims: the PTG model, the two-player wedge decomposition, the selection-robustness result, and WEDGE-CERTIFY with instance-dependent certificate complexity and matching lower bounds. These claims are then developed in the contributions and main technical sections.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The paper discusses several scope limitations: the baseline model treats transparency as commitment to a mixed action, the main wedge reduction is for two-player PTGs, and future work should study richer policy-function transparency, stochastic or continuous transparency, incomplete information, and nonstationary Stage-2 learning.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: The paper explicitly states standing assumptions SA1–SA7, including the two-player finite-game assumption, simultaneous-profile-consistent selection, bounded payoff feedback, sub-Gaussian profile-pull noise, confidence parameter conditions, and computational hypotheses. It also provides proof sketches in the main text and full proofs in the appendix

   Guidelines:

   - The answer NA means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: The paper is primarily theoretical and does not report empirical experiments. The learning/certification result is presented as an algorithmic and theoretical sample-complexity guarantee rather than an experimental evaluation.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [N/A]

   Justification: The paper does not introduce empirical datasets or experimental code required to reproduce numerical results. The main reproducibility requirements are theoretical definitions, algorithms, assumptions, and proofs.

   Guidelines:

   - The answer NA means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [N/A]

   Justification: The paper does not include training/test experiments, data splits, hyperparameters, optimizers, or empirical benchmark settings. Its results are theoretical characterizations, certification bounds, and illustrative game-theoretic examples.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

Justification: The paper does not report empirical experiments requiring error bars, confidence intervals, or statistical significance testing. The probability guarantees appear in theoretical certification bounds instead.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: Since the paper does not conduct computational experiments, there are no required CPU/GPU resources, memory specifications, runtime measurements, or total compute estimates to report.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: To the best of our knowledge, the research conforms to the NeurIPS Code of Ethics. The paper is theoretical, does not use human subjects, private data, sensitive datasets, or high-risk deployed models, and preserves anonymity in the submission.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.

- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [No]

    Justification: The paper is foundational theoretical work on equilibrium structure and learning/certification in Policy Transparency Games. It does not introduce a deployable technology, dataset, or practical automated-pricing system, and the motivating examples are illustrative rather than prescriptive. Any societal implications would be indirect and application-dependent, so a dedicated broader-impact discussion is not applicable to the work performed in this paper.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
    - If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The paper does not release high-risk models, pretrained language models, image generators, scraped datasets, or other artifacts requiring controlled release safeguards.

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

    Justification: The paper does not appear to use external datasets, pretrained models, software packages, or other existing assets requiring license documentation. The cited works are used as scholarly references only.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: The paper does not release new datasets, models, benchmark suites, or code assets. Its main contributions are theoretical definitions, structural theorems, and certification algorithms.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.

- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve human subjects, so IRB approval or equivalent ethical review is not applicable.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components.

    Guidelines:

    - The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.