Stability of Cooperation in Imperfect Conditions

The Folk Theorem with Costly, Imperfect Monitoring

Soham Pattnaik
sopatt97@gmail.com

Euler Circle

July 9, 2026

Outline

Game Theory

The One-Shot Prisoner's Dilemma

Formally define the stage game as the tuple $\mathcal{G} = (N, (A_i)_{i \in N}, (u_i)_{i \in N})$:

- **Players**: $N = \{1, 2\}$.
- **Action Spaces**: $A_1 = A_2 = \{C, D\}$, where C represents Cooperate and D represents Defect.
- **Joint Strategy Space**: $A = A_1 \times A_2 = \{(C, C), (C, D), (D, C), (D, D)\}$.
- **Payoff Mappings**: $u_i : A \to \mathbb{R}$, illustrated in the matrix below:

	Cooperate (C)	Defect (D)
Cooperate (C)	$(u_1(C, C) = 3,\ u_2(C, C) = 3)$	$(u_1(C, D) = 0,\ u_2(C, D)$
Defect (D)	$(u_1(D, C) = 5,\ u_2(D, C) = 0)$	$(u_1(D, D) = 1,\ u_2(D, D)$

Table: Stage Game Payoff Matrix $\mathcal{G}$

A strategy profile $a^* \in A$ is a Nash Equilibrium if $u_i(a_i^*, a_{-i}^*) \geq u_i(a_i, a_{-i}^*)$ for all $a_i \in A_i$. Here, the unique equilibrium is (D, D), yielding the Pareto-inefficient profile $(1, 1)$.

Repeated Games

Repeated Games & Discounted Payoffs

The stage game $\mathcal{G} = (N, (A_i)_{i \in N}, (u_i)_{i \in N})$ is played at periods $t = 1, 2, \ldots$. Players evaluate infinite streams of payoffs using the normalized discounted sum:

$$U_i = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i(a^t)$$

Where each mathematical component represents:

- $\delta \in (0, 1)$: The discount factor, reflecting player patience and game longevity.
- $1 - \delta$: The normalizing factor, scaling repeated payoffs to stage-game units.
- $a^t = (a_1^t, \ldots, a_n^t) \in A$: The joint action profile chosen at time t.
- $u_i(a^t)$: The stage-game utility achieved by player i in period t.

One-Shot vs Repeated

- One-shot PD: defect is dominant $\rightarrow$ (1,1)
- Repeated PD: future punishment can sustain cooperation (3,3)
- Key idea: "Shadow of the future"

The Folk Theorem

The Folk Theorem (Fudenberg & Maskin, 1986)

Let $\underline{v}_i = \min_{a_{-i}} \max_{a_i} u_i(a_i, a_{-i})$ be player i's pure-strategy minimax payoff. Let V be the convex hull of feasible stage payoffs.

The Folk Theorem (With Perfect Monitoring)

If V has a dimension equal to the number of players ($|N|$), then for any feasible, strictly individually rational payoff vector $v \in V$ (where $v_i > \underline{v}_i$ for all i), there exists a threshold $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$, v can be sustained as a Subgame Perfect Equilibrium (SPE) payoff vector.

Fudenberg & Maskin (1986)

Consider target payoff $v \in V$ and minimax profiles m^j targeting player j. By full dimensionality, choose a vector $\epsilon > 0$ such that $v_i - \epsilon_i > \underline{v}_i$ and punishers are rewarded downstream.

- **Phase 1 (Cooperation)**: Play the target profile generating v. If player j deviates, transition immediately to Phase 2.
- **Phase 2 (Punishment)**: Play m^j for T periods. If any player k fails to punish, restart Phase 2 targeting player k. After T periods, move to Phase 3.
- **Phase 3 (Reward)**: Play a profile yielding $v_i - \epsilon_i + \delta\eta_i$ to reward those who punished.

Incentive Compatibility Bound for δ: A potential deviator during Phase 1 balances immediate gain M against future punishment:

$$(1-\delta)M + \delta\left[(1-\delta^T)\underline{v}_i + \delta^T(v_i - \epsilon_i)\right] \leq v_i$$

As $\delta \to 1$, the loss from the permanent downshift to the reward phase dominates, sustaining cooperation.

Imperfect & Costly Monitoring

The Challenge: Imperfect Monitoring

- Real-world monitoring is noisy or private.
- Introduces **variance** in signals.
- False alarms and missed defections destabilize cooperation.

Endogenous Monitoring & Signal Structure

Let players choose monitoring efforts $m_i \in [0, 1)$ simultaneously with actions.

- **Signal Precision Structure**: Player i receives a private signal $s_i^t = a_j^t + \epsilon_i^t$. The noise is distributed as $\epsilon_i^t \sim \mathcal{N}(0, \sigma_i^2(m_i))$, where the variance strictly decreases in monitoring intensity:

$$\sigma_i^2(m_i) = \frac{\sigma_0^2}{1 - m_i}$$

- **Information Cost Function**: To match standard economic modeling frameworks, the cost function $c_i(m_i)$ is strictly increasing, twice differentiable, and strictly convex:

$$c_i(m_i) = \frac{\gamma \cdot m_i}{1 - m_i}, \quad \gamma > 0$$

  Note that $\lim_{m_i \to 1} c_i(m_i) = \infty$, ensuring perfect monitoring is infinitely expensive, preventing trivial realizations of perfect information.

Our Model

Formalization of the Noisy Dynamic Model

The model is an infinitely repeated game with private, endogenous monitoring.

- **Stage Action Space**: $X_i = A_i \times [0, 1)$, where player i chooses action $a_i \in \{$"cooperate", "defect"$\}$ and monitoring effort m_i.
- **Private Histories**: At period t, player i's accumulated private history is:
$$h_i^t = \left(a_i^1, m_i^1, s_i^1, \dots, a_i^{t-1}, m_i^{t-1}, s_i^{t-1}\right) \in H_i^t$$
- **Strategy Mapping**: A pure strategy $\sigma_i = (\sigma_i^a, \sigma_i^m)$ is a sequence of mappings:
$$\sigma_i^a : H_i^t \to A_i, \quad \sigma_i^m : H_i^t \to [0, 1)$$
- **Equilibrium Concept**: Because private histories are not mutually observed, we evaluate **Sequential Equilibria** or construct public-strategy components to establish stable **Perfect Public Equilibria (PPE)**.

Theoretical Results

Extension of the Folk Theorem to Costly Monitoring

To establish the Folk Theorem under endogenous monitoring, we utilize the APS framework (Abreu, Pearce, Stacchetti 1990) for self-generating payoff sets. Let $V(c, \sigma^2)$ be the set of sustainable PPE payoffs.

Theorem (Folk Theorem with Costly Endogenous Monitoring)

Suppose there exists a monitoring effort $m^* \in [0, 1)$ such that the induced signal structure satisfies the pairwise identifiability condition. Then, as $\delta \to 1$, any feasible, strictly individually rational payoff vector v satisfying:

$$v_i > \underline{v}_i + c_i(m^*)$$

can be sustained as a Perfect Public Equilibrium payoff vector.

The set of enforceable payoffs strictly contracts as baseline noise σ_0^2 increases or monitoring cost efficiency γ deteriorates, creating an interior bound on viable cooperation.

Mathematical Analysis: Enforcing Cooperation

Consider a symmetric trigger strategy profile where players choose $m_i = m^*$ and cooperate (C) on the equilibrium path. If a noisy signal s_i indicates defection with a likelihood ratio exceeding threshold τ, players shift to a trigger punishment state.

Let p_{CC} be the probability of a false alarm (punishing despite cooperation) and p_{DC} be the detection probability under defection. For cooperation to be sustained, the incentive compatibility constraint requires:

$$\begin{aligned}(1-\delta)[u(C,C)-c(m^*)] &+ \delta[(1-p_{CC})V_C + p_{CC}V_D] \\ &\geq (1-\delta)u(D,C) + \delta[(1-p_{DC})V_C + p_{DC}V_D]\end{aligned}$$

Rearranging terms yields the analytical **patience threshold bound**:

$$\delta \geq \frac{u(D,C)-u(C,C)+c(m^*)}{u(D,C)-u(C,C)+c(m^*)+(p_{DC}-p_{CC})(V_C-V_D)}$$

This explicitly demonstrates that higher monitoring efforts m^* maximize the precision spread $(p_{DC}-p_{CC})$, neutralizing temptation even when δ is bounded away from 1.

Conditions for Robustness

- High patience (δ large)
- Credible punishments
- Incentive-compatible monitoring investment

Simple grim-trigger adaptations work well when these conditions hold.

Computational Model

Evolutionary Simulation

- Three player types: Fixed cost, Adaptive, None (Blind).
- Population evolves based on success (fitness = total points).
- Mutation and selection on parameters (cost, wariness, forgiveness).
- Multiple matches per generation.

Evolutionary Architecture: State Space

We formalize our agent-based computational script as a dynamic evolutionary system $\mathcal{E} = (P^g, \mathcal{W}_i, \mathcal{S}, \mathcal{M})$:

Population Configuration (P^g)

- **Total Capacity**: $M = 150$ agents per generation g.
- **Uniform Seeding**: Initialized evenly with $\Delta M = 50$ per type.
- **Trait Vectors**: $\theta_i = (\text{type}, \bar{c}_i, \omega_i, \phi_i)$ track strategy genes.

Stochastic Fitness Assignment ($\mathcal{W}_i$)

Each generation, agents play $K = 5$ random matches for $T = 100$ rounds:

$$\mathcal{W}_i = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{T} \sum_{t=1}^{T} \left[u_i(a_i^t, a_k^t) - c_i^t(m_i^t) \right]$$

Fitness is strictly net utility minus monitoring cost points.

Evolutionary Architecture: Strategy Operators

The system updates parameters across generations via selection and genetic noise channels:

- **Truncation Selection ($\mathcal{S}$)**: Sorts agents by $\mathcal{W}_i$.
- **Culling Rate**: Bottom 50% (75 agents) are instantly deleted.
- **Replication**: Top 50% survive and clone replacement offspring.

Stochastic Parametric Mutation ($\mathcal{M}$)

Cloned offspring traits face random walk perturbations if probability conditions pass:

- **Base Cost Ceiling**: $\bar{c}_{\text{child}} \sim \mathcal{N}(\bar{c}_{\text{parent}}, 0.10^2)$ (prob: 0.22, lower bound: 0.01)
- **Wariness Factor**: $\omega_{\text{child}} \sim \mathcal{N}(\omega_{\text{parent}}, 0.03^2)$ (prob: 0.18, range: $[0.01, 0.5]$)
- **Forgiveness Rate**: $\phi_{\text{child}} \sim \mathcal{N}(\phi_{\text{parent}}, 0.03^2)$ (prob: 0.18, range: $[0.0, 0.5]$)

Simulation Results: Visualizations

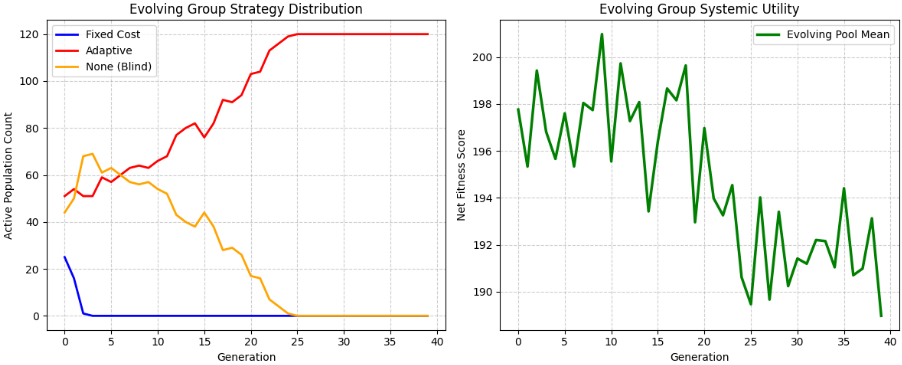

Figure: Simulation Data

Simulation Results: Dominant Agents

Running the code across $G = 40$ generations reveals survival trends:

- **Adaptive Dominance**: The Adaptive agents were able to fill the entire population space.

Failures of Alternate Strategies

Fixed-Cost Profile Likely died off due to the evolutionary pressure being too great. If the initial variables were perfect, it may have done better, but still would have lost to the Adaptive agents due to the inability to build trust.

Blind Profile ($m_i = 0$) High ambient noise ($\sigma_0^2 = 5$) creates essentially blindness to the opponents' behavior. Victims are ruthlessly exploited by defectors. The only reason they lasted so long is because of the lack of monitoring cost.

Simulation Results: Trait Convergence

The evolutionary optimization tracking uncovers highly specific mathematical trait boundaries necessary to realize the Folk Theorem under noisy private conditions ($M_{\text{adaptive}} = 120$):

- **Base Cost Ceiling ($\bar{c}_i$)**: Settles at $\mu = 0.4212$ ($\sigma = 0.0414$).
- **Systemic Wariness (ω_i)**: Settles at $\mu = 0.1600$ ($\sigma = 0.0278$).
- **Stochastic Forgiveness (ϕ_i)**: Settles at $\mu = 0.0176$ ($\sigma = 0.0154$).

Trait Blueprint of the Highest Performing Elite Agent

The single most successful strategy profile discovered by selection pressure uses:

$$\theta_{\text{elite}} = (\bar{c}^* = 0.4121,\ \omega^* = 0.1683,\ \phi^* = 0.0268)$$

Simulation Results: Trait Convergence (cont'd)

Strategic Mechanics of the Converged Blueprint

- **Wariness**: A wariness scalar of ~ 0.16 prevents immediate trust collapse over single noise spikes while remaining sensitive to systemic defection.
- **Forgiveness**: A low but nonzero forgiveness value ($\sim 1.7\%$) ensures that error correction exists but minimizes vulnerability to probers.

Conclusion

Implications

- Collusion with costly auditing
- Partnerships and oversight
- Institutional design (regulation, tax agencies)

Main Takeaway

Even with noisy and costly monitoring, sufficiently patient players can often sustain near-efficient cooperation by carefully balancing monitoring expenditures against the gains from cooperation.

Future work: network monitoring, continuous time, learning dynamics.

Thank You!

Questions?