A Game-Theoretic Interpretation of Organized Crime and Institutions

Andrés Romero

Players and Strategies

Two players interact in a simultaneous-move game:

- The State $E \in \{F, D\}$ chooses **Strong Institutions** (F) or **Weak Institutions** (D).
- Organized Crime $O \in \{N, X\}$ chooses **Non-confrontation** (N) or **Confrontation/Expansion** (X).

Parameters

Non–negative parameters:

- k: cost of maintaining F (enforcement, monitoring).
- λ: legitimacy/efficiency premium for the State under (F, N).
- D: institutional damage to the State if (D, X).
- d: analogous damage under (F, X) ($d \leq D$).
- $\rho \in [0, 1]$: effectiveness of repression under F.
- m: extra operational cost for F under confrontation.
- r_F, r_D: O's rents under non-confrontation with F or D ($r_D \geq r_F$).
- R_F, R_D: O's rents under confrontation with F or D ($R_D \geq R_F$).
- c: organizational cost of confrontation for O.
- s: sanction if repressed; under F the effective sanction is $s\rho$.
- $\theta \geq 0$: legitimacy loss from choosing D even without confrontation.

Payoff Matrix

	N	X
F	$\big(\lambda - k,\ r_F\big)$	$\big(-k - m - d(1-\rho),\ R_F - c - s\rho\big)$
D	$\big(-\theta,\ r_D\big)$	$\big(-\theta - D,\ R_D - c\big)$

First entry: State; second: Organized Crime.

Best Responses (summary)

Organized Crime O:

$$\begin{aligned} \text{Under } F: \quad & X \succ N \iff R_F - r_F \geq c + s\rho, \\ \text{Under } D: \quad & X \succ N \iff R_D - r_D \geq c. \end{aligned}$$

State E:

$$\begin{aligned} \text{Given } N: \quad & F \succ D \iff \lambda \geq k - \theta, \\ \text{Given } X: \quad & F \succ D \iff D - d(1-\rho) - m + \theta \geq k. \end{aligned}$$

Pure Strategy Nash Equilibria

- (F, N) iff $\lambda \geq k - \theta$ and $R_F - r_F \leq c + s\rho$.
- (D, N) iff $\lambda < k - \theta$ and $R_D - r_D < c$.
- (F, X) iff $D - d(1-\rho) - m + \theta \geq k$ and $R_F - r_F > c + s\rho$.
- (D, X) iff $D - d(1-\rho) - m + \theta < k$ and $R_D - r_D \geq c$.

Mixed Strategy Equilibrium (summary)

Let p be the probability that O plays X, and q the probability that E plays F. Indifference yields

$$p = \frac{k - \theta - \lambda}{D - d(1-\rho) - m - \lambda}, \qquad q = \frac{c - (R_D - r_D)}{(R_F - R_D) - (r_F - r_D) - s\rho}.$$

A mixed equilibrium exists if $p, q \in (0, 1)$ and denominators are nonzero.

Comparative Statics (intuition)

- Higher $\rho \Rightarrow$ larger $s\rho$ under F; X becomes less attractive; (F, N) expands.
- Higher $k \Rightarrow F$ becomes costlier; $(F, \cdot)$ regions shrink; p rises in the mixed equilibrium.
- Higher D or lower $d \Rightarrow$ strengthens F when confrontation is likely; shifts mass from (D, X) toward (F, X) or (F, N).
- Larger $R_D - r_D \Rightarrow$ increases the likelihood of (D, X) (collapse zone).

Interpretation

1. With effective institutions, governance coordinates on low confrontation (F, N).
2. With weak institutions, low confrontation (D, N) prevails until $R_D - r_D \geq c$, after which confrontation dominates.
3. Collapse arises under (D, X) when the State avoids F and criminal returns to confrontation exceed costs.

Signalling Extension (Perfect Bayesian Equilibrium)

Nature first draws the State's type: **Strong** (S) with probability π, or **Weak** (W) with probability $1-\pi$. The State observes its type and chooses a signal $s \in \{s_{strong}, s_{weak}\}$. Organized Crime O observes s (but not the type) and then chooses $a \in \{N, X\}$.

Type-dependent payoffs (mapping to baseline)

When O chooses N:

$$U_E^S(N) = \lambda - k, \quad U_E^W(N) = -\theta; \qquad U_O^S(N) = r_F, \quad U_O^W(N) = r_D.$$

When O chooses X:

$$U_E^S(X) = -k-m-d(1-\rho), \quad U_E^W(X) = -\theta-D; \quad U_O^S(X) = R_F-c-s\rho, \quad U_O^W(X) = R_D-c.$$

Beliefs and O's best response to a signal

Let $\mu \equiv \Pr(S \mid s)$ be O's belief after seeing s. The expected gain from X over N is

$$\Delta_O(\mu) = \mu\big[(R_F - r_F) - (c + s\rho)\big] + (1-\mu)\big[(R_D - r_D) - c\big].$$

O chooses X iff $\Delta_O(\mu) \geq 0$. Define the belief threshold $\mu^\star$ by $\Delta_O(\mu^\star) = 0$:

$$\boxed{\mu^\star \;=\; \frac{c-(R_D-r_D)}{(R_F-R_D)-(r_F-r_D)-s\rho}}.$$

Hence O plays X after s iff $\mu \geq \mu^\star$ (given the denominator > 0).

Separating PBE (deterrence via strong signal)

Consider a separating profile where S sends s_{strong} and W sends s_{weak}; O's beliefs are $\mu(s_{strong}) = 1$, $\mu(s_{weak}) = 0$. O's best replies:

$$\text{after } s_{strong}: \ \Delta_O(1) = (R_F - r_F) - (c + s\rho) \overset{!}{<} 0 \Rightarrow a = N; \qquad \text{after } s_{weak}: \ \Delta_O(0) = (R_D - r_D) - c \overset{!}{\geq} 0 \Rightarrow a =$$

Let the cost of sending s_{strong} be τ_S for type S and τ_W for type W; normalize the cost of s_{weak} to zero for both types. Incentive compatibility (IC):

$$\text{IC-S:} \quad (\lambda - k) - \tau_S \;\geq\; -k - m - d(1-\rho), \;\iff\; \boxed{\tau_S \;\leq\; \lambda + m + d(1-\rho)};$$

$$\text{IC-W:} \quad (-\theta - D) - 0 \;\geq\; (-\theta) - \tau_W, \;\iff\; \boxed{\tau_W \;\geq\; D}.$$

Thus a separating PBE with deterrence exists if

$$\boxed{R_F - r_F < c + s\rho, \quad R_D - r_D \geq c, \quad \tau_S \leq \lambda + m + d(1-\rho), \quad \tau_W \geq D.}$$

Pooling on strong signal

Suppose both types send s_{strong}, so $\mu(s_{strong}) = \pi$. O chooses N after s_{strong} iff $\Delta_O(\pi) < 0$. IC for types (no profitable deviation to s_{weak} which would induce X):

$$\text{IC-S (no dev.):} \quad (\lambda - k) - \tau_S \;\geq\; -k - m - d(1-\rho) \iff \tau_S \leq \lambda + m + d(1-\rho),$$

$$\text{IC-W (no dev.):} \quad (-\theta) - \tau_W \;\geq\; (-\theta - D) \iff \tau_W \leq D.$$

Pooling is sustainable when

$$\boxed{\Delta_O(\pi) < 0, \qquad \tau_S \leq \lambda + m + d(1-\rho), \qquad \tau_W \leq D}.$$

Note the complementarity with separation: separation requires $\tau_W \geq D$, while pooling requires $\tau_W \leq D$.

Discussion

Signalling rationalizes bluffing: weak states may mimic strong signals when τ_W is low, leading to pooling and potential "tests" by O. When strong signals are sufficiently costly for weak states ($\tau_W \geq D$), only strong states can credibly send them, producing separation and deterrence.

Appendix A. Step-by-step derivations (simultaneous game)

A1. State's best response given N

$$u_E(F,N) = \lambda - k, \qquad u_E(D,N) = -\theta.$$

Prefer F iff

$$\lambda - k \geq -\theta \iff \boxed{\lambda \geq k - \theta}.$$

A2. State's best response given X

$$u_E(F,X) = -k - m - d(1-\rho), \qquad u_E(D,X) = -\theta - D.$$

Prefer F iff

$$\begin{aligned} -k - m - d(1-\rho) &\geq -\theta - D \\ \theta + D - m - d(1-\rho) &\geq k \\ \iff \quad \boxed{D - d(1-\rho) - m + \theta \geq k}. \end{aligned}$$

A3. O's best response under F

$$u_O(F,N) = r_F, \qquad u_O(F,X) = R_F - c - s\rho.$$

Prefer X iff

$$\boxed{R_F - r_F \geq c + s\rho}.$$

A4. O's best response under D

$$u_O(D,N) = r_D, \qquad u_O(D,X) = R_D - c.$$

Prefer X iff

$$\boxed{R_D - r_D \geq c}.$$

A5. Mixed equilibrium: State indifferent (solve for p)

$\mathbb{E}[u_E|F] = (1-p)(\lambda - k) + p\big[-k - m - d(1-\rho)\big], \quad \mathbb{E}[u_E|D] = (1-p)(-\theta) + p(-\theta - D).$

Indifference $\Rightarrow$

$$\begin{aligned} (1-p)(\lambda - k) + p[-k - m - d(1-\rho)] &= (1-p)(-\theta) + p(-\theta - D) \\ p\,[D - d(1-\rho) - m - \lambda] &= k - \theta - \lambda \\ \implies \quad \boxed{p = \frac{k - \theta - \lambda}{D - d(1-\rho) - m - \lambda}}. \end{aligned}$$

A6. Mixed equilibrium: O indifferent (solve for q)

$$\mathbb{E}[u_O|X] = q(R_F - c - s\rho) + (1-q)(R_D - c), \quad \mathbb{E}[u_O|N] = q\, r_F + (1-q) r_D.$$

Indifference $\Rightarrow$

$$q\big[(R_F - R_D) - s\rho - (r_F - r_D)\big] = c - (R_D - r_D)$$

$$\Longrightarrow \quad \boxed{q = \frac{c - (R_D - r_D)}{(R_F - R_D) - (r_F - r_D) - s\rho}}.$$

Equivalent form (multiply num/den by -1):

$$\boxed{q = \frac{(R_D - r_D) - c}{(R_D - R_F) - (r_D - r_F) + s\rho}}.$$

A7. Interior-existence checks

A mixed equilibrium with $p, q \in (0,1)$ exists when denominators are nonzero and parameter values imply $0 < p < 1$, $0 < q < 1$. Otherwise, one of the pure equilibria obtains.