

# Experimental Design

## Part I: Do LLMs Play the Game?

### Pure Global Game

Private signals only  
Simultaneous decisions

### Communication

Pre-play messaging  
(Watts-Strogatz network)

### Scramble

Briefings randomly  
redistributed across periods

### Flip

Z-score negated  
before briefing

**Key findings:**

- Pure: mean  $r = +0.73$
- Scramble:  $r \rightarrow +0.23$
- Flip:  $r \rightarrow -0.67$
- Comm: +3.7 pp

## Part II: Information Design

### Stability Design

4x clarity width  
Flatter direction slope

### Instability Design

0.15x clarity width  
Steeper direction slope

### Public Signal

Shared news bulletin  
from  $\theta$

### Censorship

Upper / lower  
binary signal pooling

### Surveillance

"Communications  
monitored" warning

### Propaganda

$k$  regime plants  
send pro-regime msgs

*Falsification  
tests*

Shared: 9 models | 25 agents/period |  $\sigma = 0.3$  | temp = 0.7 | narrative briefings, no payoff tables

**Key findings:**

- Stability: +19.5 pp
- Public signal: -10.7 pp
- Censorship: +18.5 pp
- Surveillance: -17.5 pp
- Surv + censor: 30.9%  $\rightarrow$  3.7%