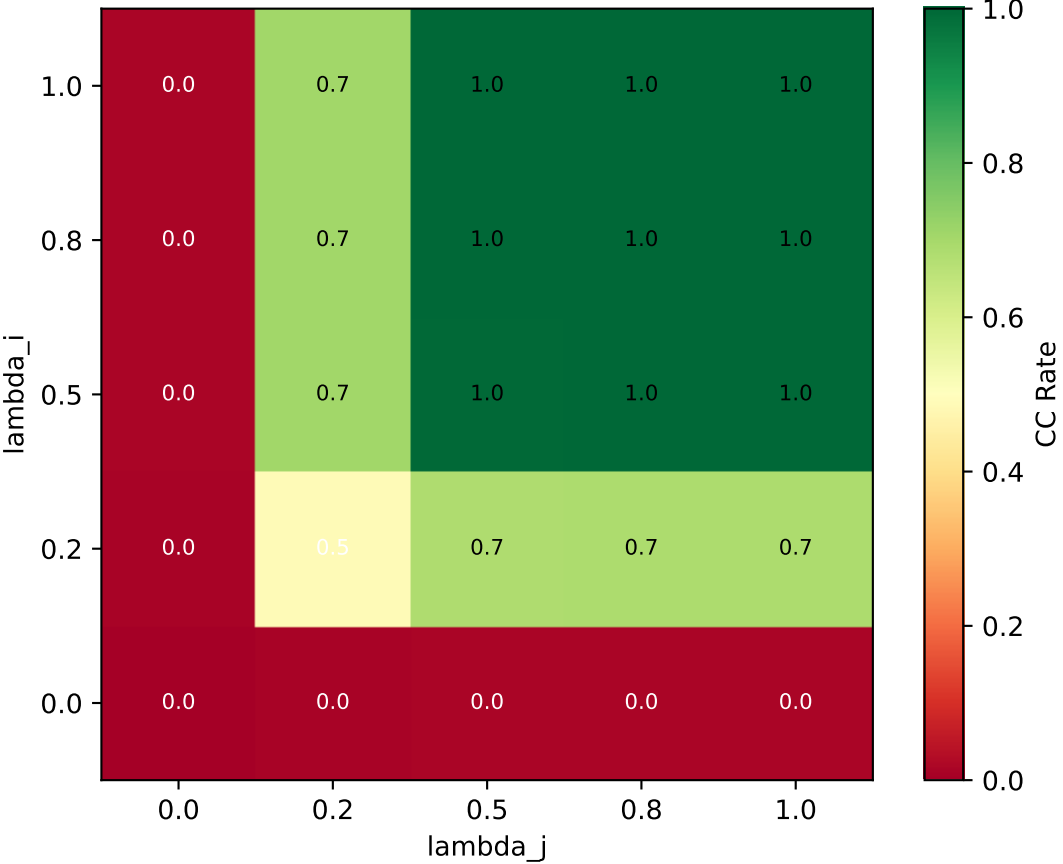
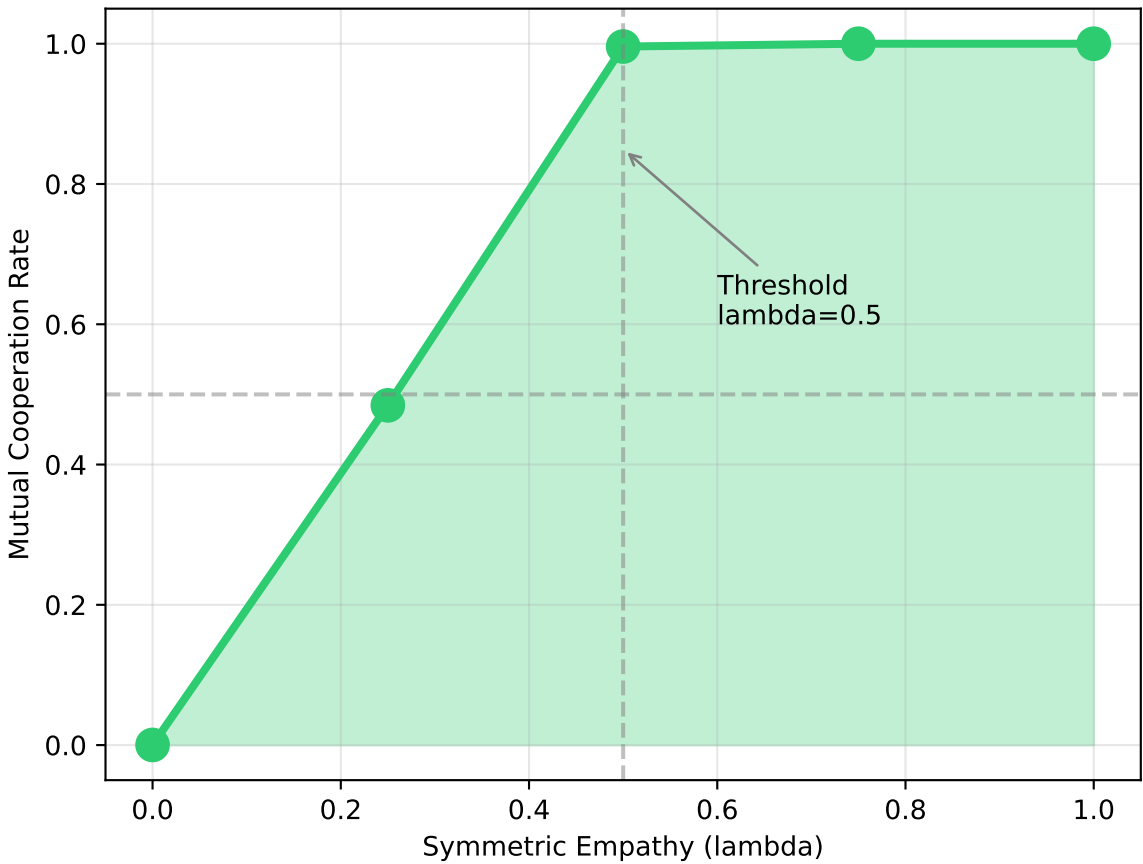


# Empathy + ToM in Prisoner's Dilemma: Phase 5 Results

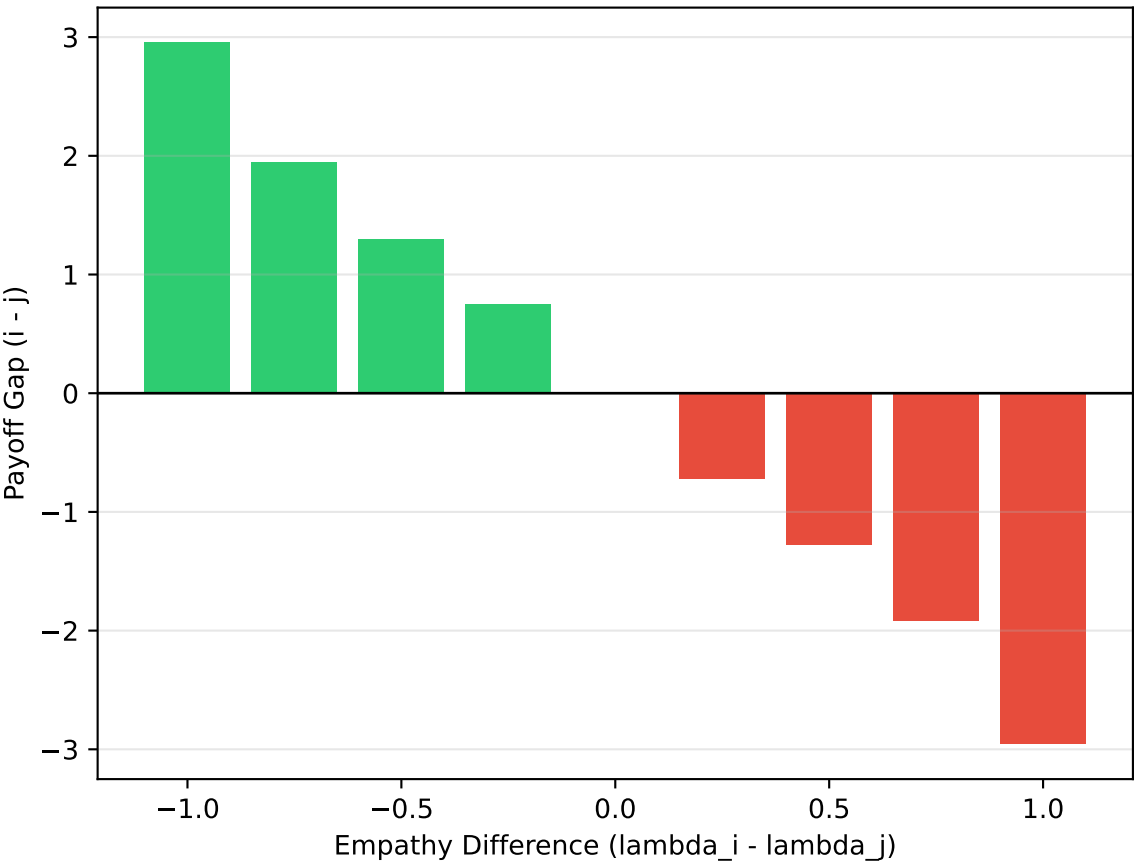
A. Cooperation Rate by Empathy



B. Cooperation Threshold Effect



C. Exploitation: High Empathy Disadvantaged



D. Summary Statistics

## D. Key Findings

Total experiments: 22,500

High empathy (both  $\geq 0.5$ ):  
- Cooperation rate: 99.9%  
- N = 8,100

Low empathy (both  $\leq 0.25$ ):  
- Cooperation rate: 12.6%  
- N = 3,600

Asymmetric ( $|\text{diff}| \geq 0.5$ ):  
- i exploited: 25.0%  
- j exploited: 25.0%  
- N = 10,800

Main Result:  
Empathy  $\geq 0.5$  reliably produces mutual cooperation without baking in pro-sociality.