Games and Risk: A Foundational Perspective

Allison C. Reilly & Seth D. Guikema
University of Michigan
May 19, 2016

Outline of Talk

1) Strategic behavior in the context of risk analysis
   - What is Risk? Rationality? Strategic Behavior?
   - Foundational Assumptions in Game Theory
2) To what extent do our assumptions about opponents matter in games?
   - Would a defender change their actions if they knew the attacker was not “rational?”
   - Does being wrong about assumptions impacted expected payoffs?
3) Concluding thoughts

What is risk?

Risk = f(Uncertainty, Consequences)

Glossary, Society for Risk Analysis

- Risk is the possibility of an unfortunate occurrence
- Risk is the potential for realization of negative consequences of an event
- Risk is exposure to a proposition (e.g. the occurrence of a loss) of which one is uncertain
- Risk is uncertainty about and severity of the consequences of an activity with respect to something that humans value

- Risk is NOT:
  - Risk ≠ Probability Risk ≠ P x C
  - Risk ≠ Consequence Risk ≠ E(U)

What is rationality?

Rationality

A decision-maker can be considered rational if they act is if they maximize a utility function, consistent with the axioms of rational choice

1. Orderability 2. Continuity 3. Substitutability 4. Monotonicity
5. Decomposability

- Does rationality imply "excessive knowledge?" No.
- Does rationality imply "strategic?" No.

What does it mean to be strategic?

Strategic

A strategic player in a given *game* is one who reasons about her opponent's actions and responds in a manner

1. consistent with her beliefs and reasoning about the opponent's behavior and
2. her own decision-making framework

Reconciling Strategic and Rational

	Strategic	Non-Strategic
Rational	Maximizes utility function and, inferring likely opponent action, chooses appropriate response	Maximizes utility function without considering opponent action
Non-Rational	Infers likely opponent action but makes decisions in a way other than utility maximization (e.g., satisficing, quantal response, etc.)	Chooses without maximizing utility or considering opponent actions

Three Key Foundational Assumptions of Game Theory

1. Instrumental rationality & common knowledge of rationality
2. Knowledge of the rules of the game
3. Common Priors

1: Instrumental Rationality

- Standard game theory assumes rational players *and* common knowledge of rationality
  - This is a central assumption to the calculation of a Nash Equilibrium
- Note:
  - This assumption means that all players are rational and everyone knows that everyone else is rational
  - This assumption **does not** mean that utility functions have to be common knowledge

2: Knowledge of the Rules of the Game

- Standard game theory assumes that everyone knows the “rules.” This includes:
  - The action set of each player
  - What each player knows when they make their choice
  - When each player makes their choice relative to when the other players make their choice
- Note:
  - Assumption disqualifies standard game theory from modeling “black swan” actions (i.e., actions by opponents that were not thought possible before they occurred)

3: Common Prior

- Standard game theory assumes that all players start with a common prior and update from this based on their experiences
  - Important in Harsanyi's reduction of a game of incomplete info. to a game of complete info
  - Aumann's recursive argument attempts to justify it
- Note:
  - Means that opponents in a game are assumed to start from the same probability distribution, and any differences are due *only* to differences in information
  - Means that fundamental differences in beliefs do not fit within the standard framework

Luce & Raiffa (1957)

"... game theory is not descriptive, but rather (conditionally) normative. It states neither how people do behave *nor how they should behave in an absolute sense*, but how they should behave if they wish to achieve certain ends. It prescribes for given assumptions courses of action for the attainment of outcomes having certain formal 'optimum' properties. These properties may or may not be deemed pertinent in any given real world conflict of interest. If they are, the theory prescribes the choices which must be made to get that optimum."

Reconciling Strategic and Rational

	Strategic	Non-Strategic
Rational	Maximizes utility function and, inferring likely opponent action, chooses appropriate response	Maximizes utility function without considering opponent action
Non-Rational	Infers likely opponent action but makes decisions in a way other than utility maximization (e.g., satisficing, quantal response, etc.)	Chooses without maximizing utility or considering opponent actions

What Does It Mean to Apply Standard Game Theory to A Risk Problem

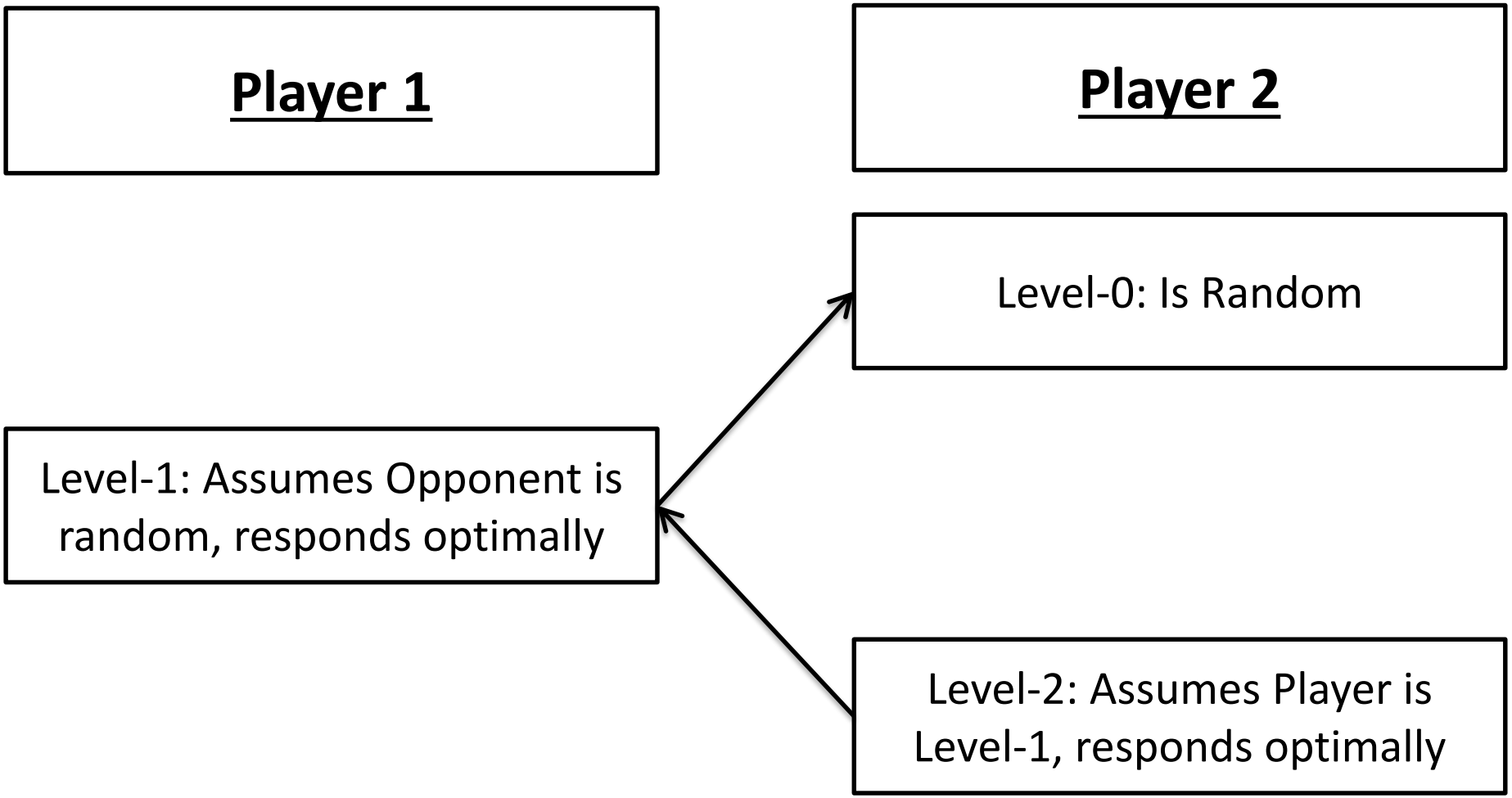
1. You must agree that the underlying assumptions are reasonable
2. You must fully characterize the beliefs and preferences of all players together with the rules of the game
3. You must address probabilistic outcomes over the consequence space in accordance with the definition of risk (expectations are insufficient)

EXPANDING THE CONVERSATION: BROADENING BEYOND STANDARD GAME THEORY

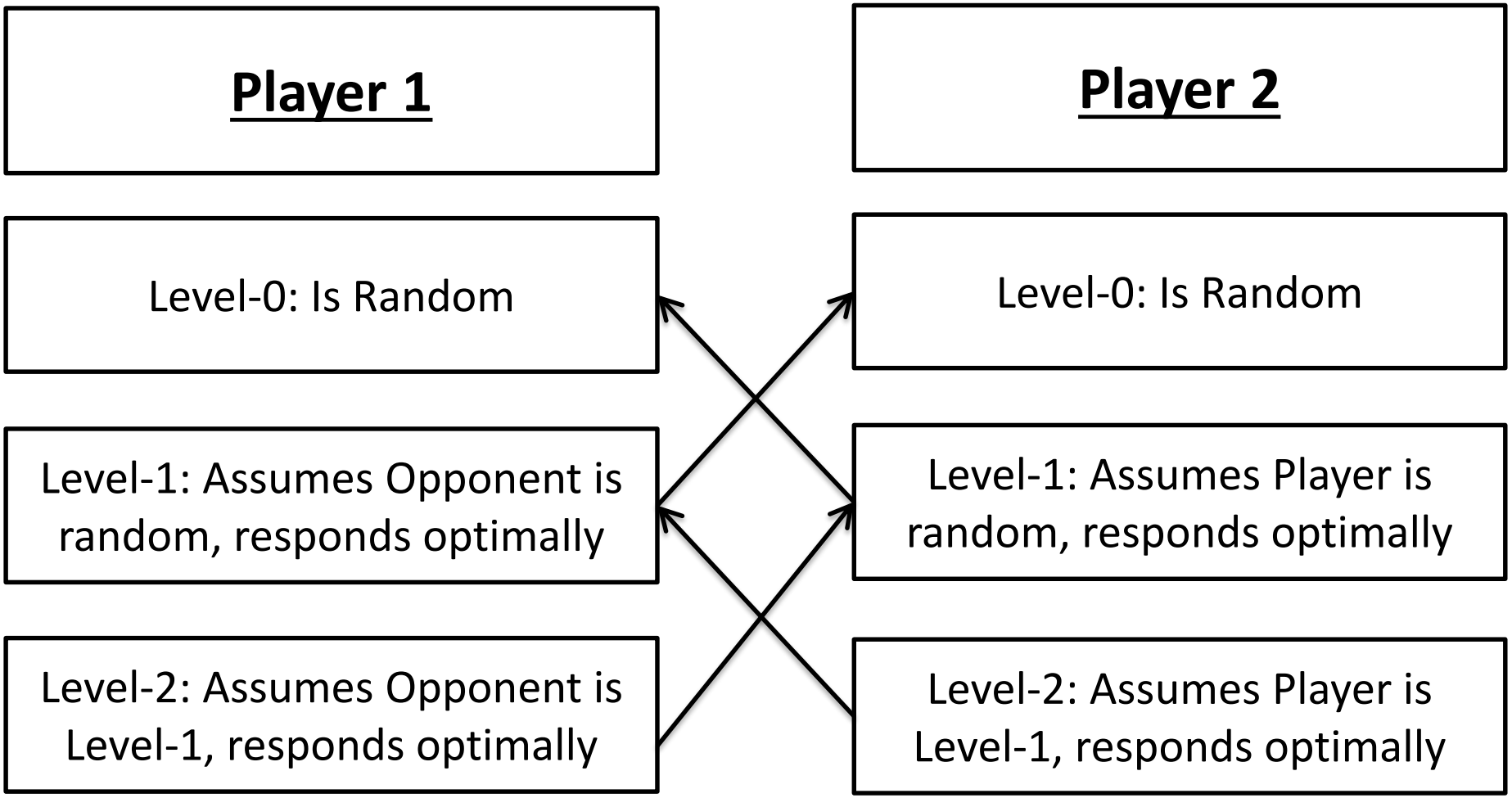
Reconciling Strategic and Rational

	Strategic	Non-Strategic
Rational	Maximizes utility function and, inferring likely opponent action, chooses appropriate response	Maximizes utility function without considering opponent action
Non-Rational	Infers likely opponent action but makes decisions in a way other than utility maximization (e.g., satisficing, quantal response, etc.)	Chooses without maximizing utility or considering opponent actions

What is level-k thinking?



What is level-k thinking?



Motivation

- “Make everyone rational. This way we will know the worst case.”
- Explore implications of foundational assumption #1 (Instrumental Rationality)
  - What are the implications of assuming a “community of rationality” when such a community does not exist?
  - What *should* we assume about opponent’s strategic reasoning?

Motivating Case 1

	Attacker Attacks	Attacker Doesn't Attack
Defender Defends	(1, 1)	(-5, 0)
Defender Doesn't Defend	(-8, 10)	(5, 0)

Outcomes

(1) Nash Equilibrium (Pure Strategy):

- [Attack/Defend] A = 1, D = 1

Motivating Case 1

	Attacker Attacks	Attacker Doesn't Attack
Defender Defends	(1, 1)	(-5, 0)
Defender Doesn't Defend	(-8, 10)	(5, 0)

Outcomes

(1) Nash Equilibrium (Pure Strategy):

- [Attack/Defend] A = 1, D = 1

(2) What if defender is wrong about attacker's decision framework?

- Defender assumes a community of rationality; Attacker is level-1 (assumes defender is random): [Attack/Defend] A = 1, D = 1

→ Suggests that assuming a community of rationality doesn't make the defender worse off.

Motivating Case 2

	Attacker Attacks	Attacker Doesn't Attack
Defender Defends	(25, 11)	(-5, 10)
Defender Doesn't Defend	(-50, 10)	(10, 10)

Outcomes

(1) Nash Equilibrium (Pure Strategy):

- [No Attack/No Defense] A = 10, D = 10

Motivating Case 2

	Attacker Attacks	Attacker Doesn't Attack
Defender Defends	(25, 11)	(-5, 10)
Defender Doesn't Defend	(-50, 10)	(10, 10)

Outcomes

(1) Nash Equilibrium (Pure Strategy):

- [No Attack/No Defense] A = 10, D = 10

(2) What if defender is wrong about attacker's decision framework?

- Defender assumes a community of rationality; Attacker is level-1 (assumes defender is random): [Attack/No Defense] A = -50, D = 10

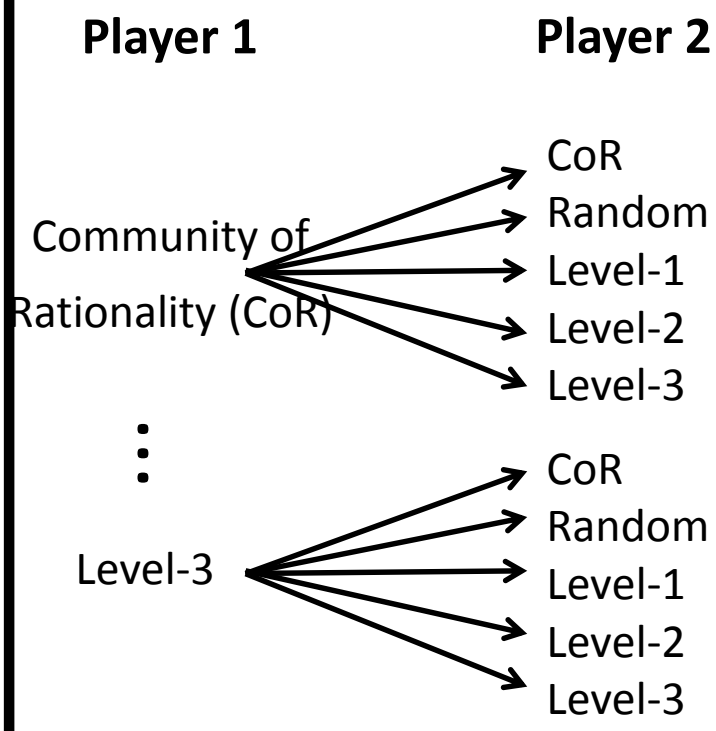
→ Assuming a community of rationality and being wrong can make the defender worse off!

Methodology

1. Simulate the interactions of 2 players via thousands of games and combinations of strategic sophistication
2. Consider how correlation between payoff matrices affect results
3. Consider both Bayesian and non-Bayesian games

Methodology

Combinations of strategic sophistication considered



Games without nature

1. Generate 1,000 3 x 3 games for a range of correlations
2. Find the expected interactions for all combinations of strategic sophistication

Metrics of evaluation

1. Best outcome most often
2. Worst outcome least often
3. Minimizes aggregated losses over range of games

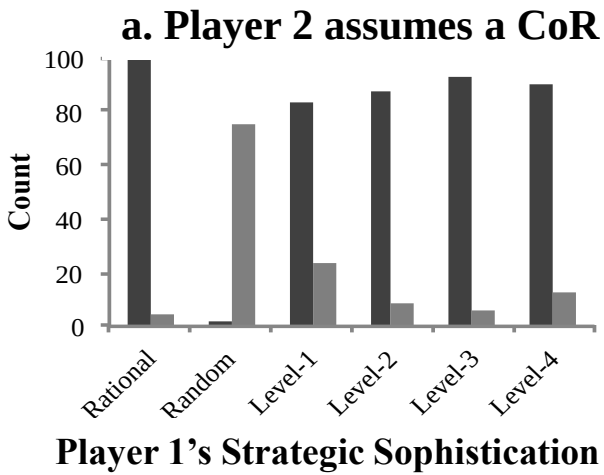
} Regret

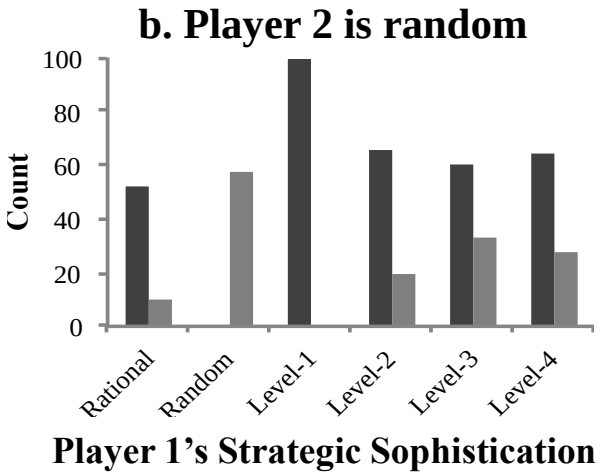
Friend or Foe?

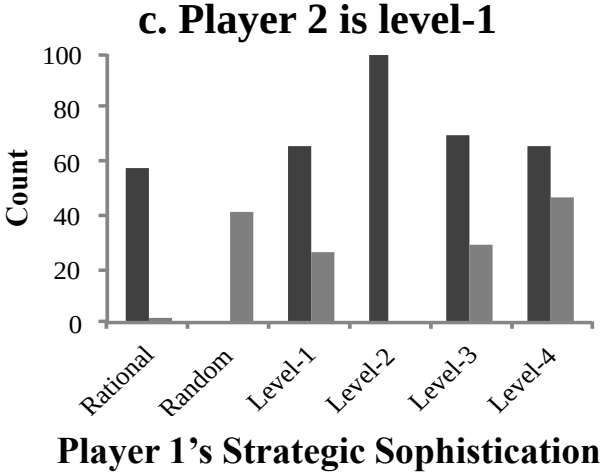
Intuitively, conservative assumptions may depend on whether our opponent is a friend or adversary

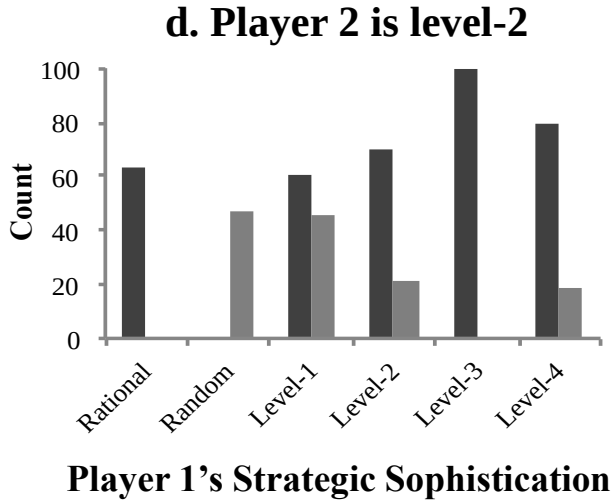
Number of times P1's assumption about P2 is best (equivalent to "correct" assumption) and worst

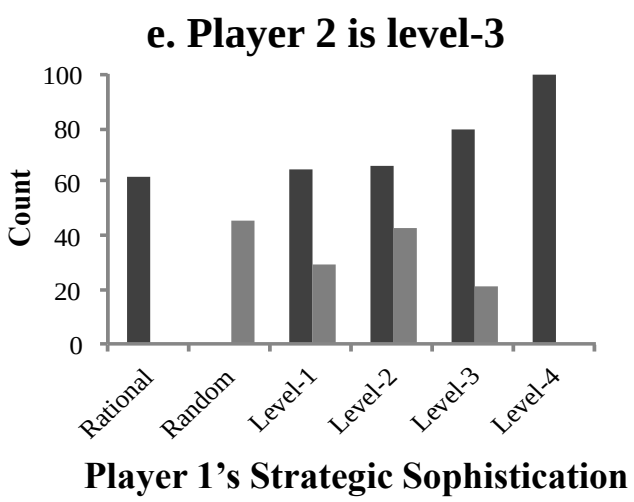
(correlation: -0.70 to -0.72)





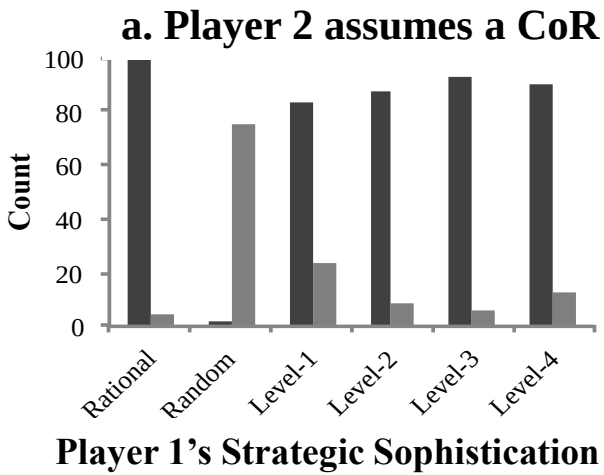


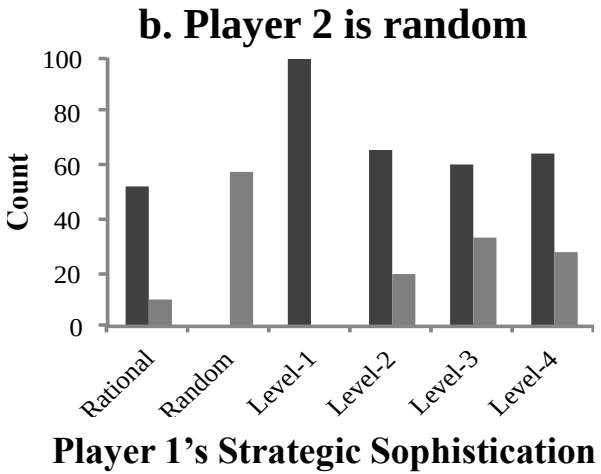


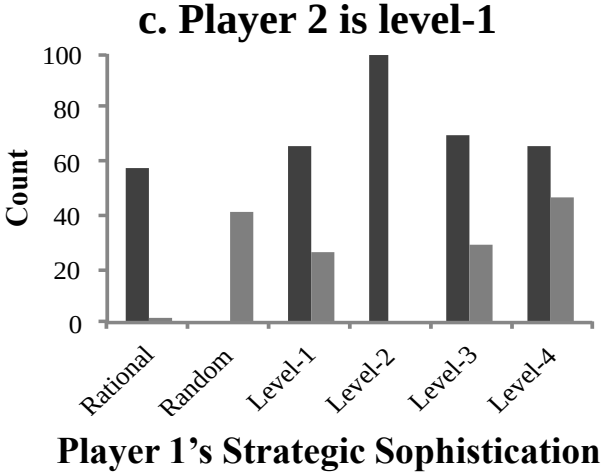


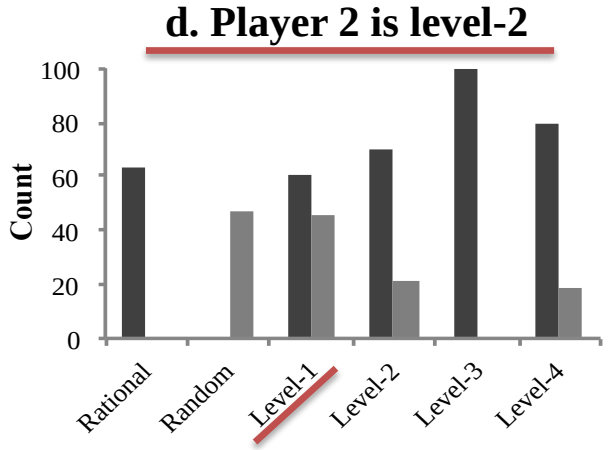
Number of times P1's assumption about P2 is best (equivalent to "correct" assumption) and worst

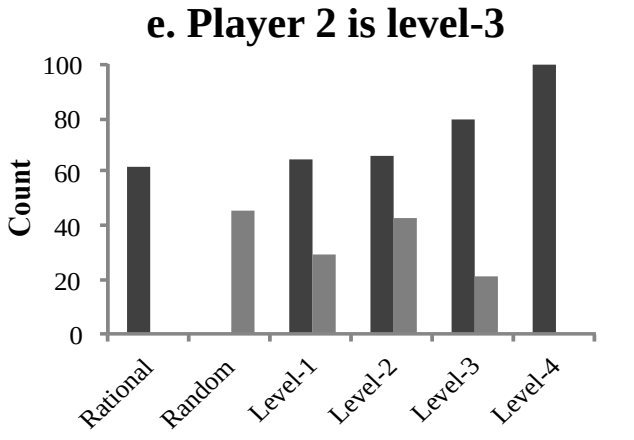
(correlation: -0.70 to -0.72)





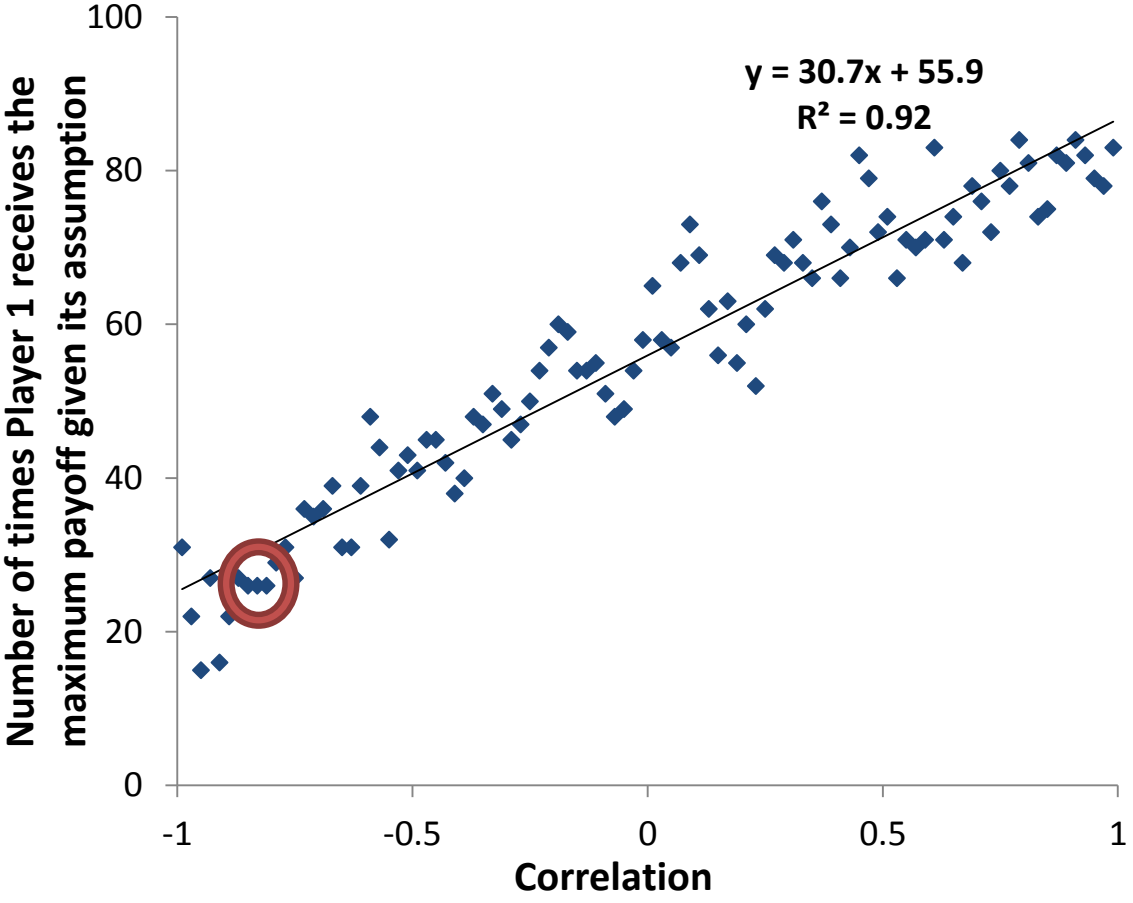




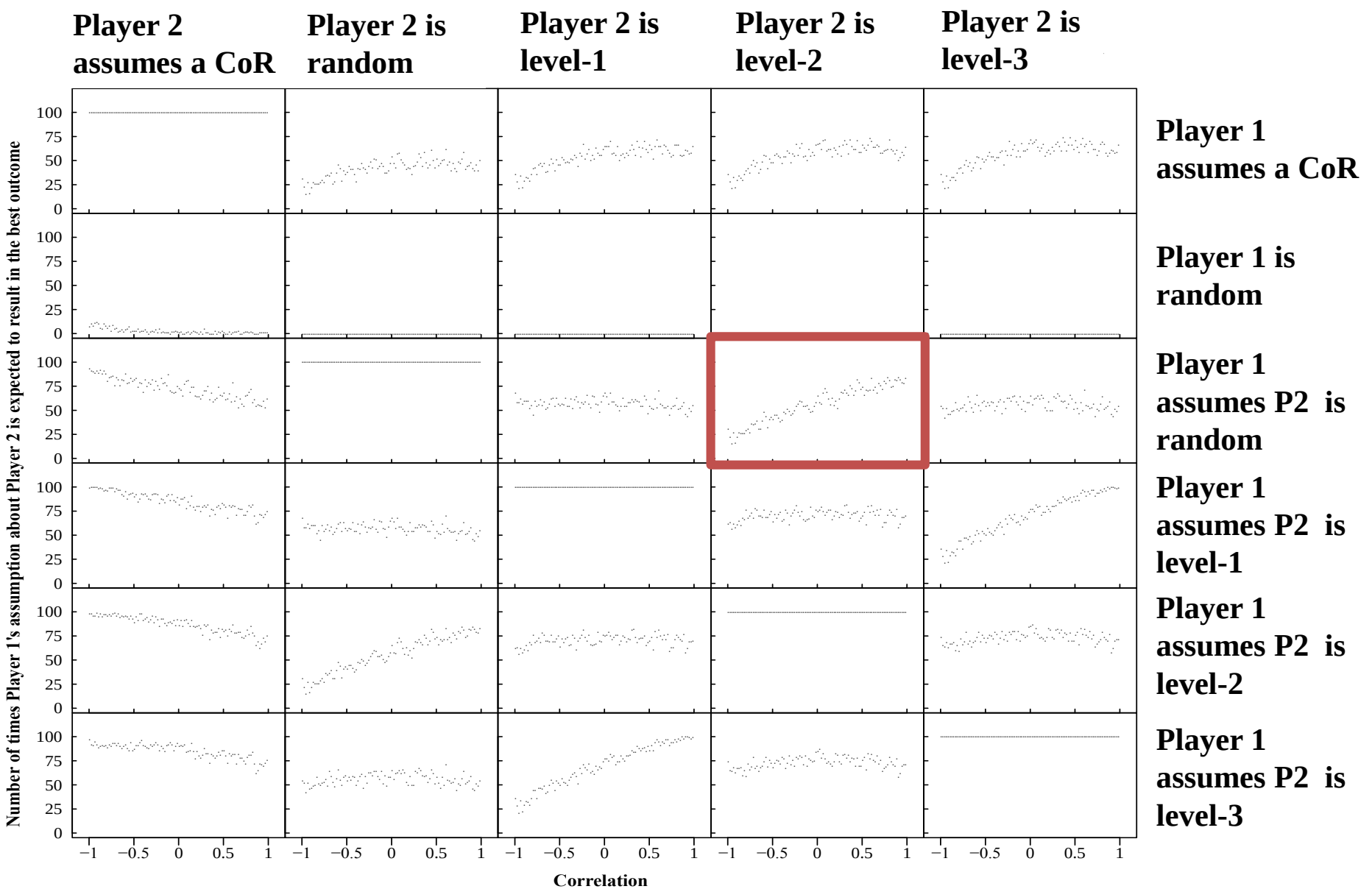


Number of times P1's assumption about P2 is best (equivalent to "correct" assumption)

Player 1 assumes that Player 2 acts randomly. Player 2 is level-2.

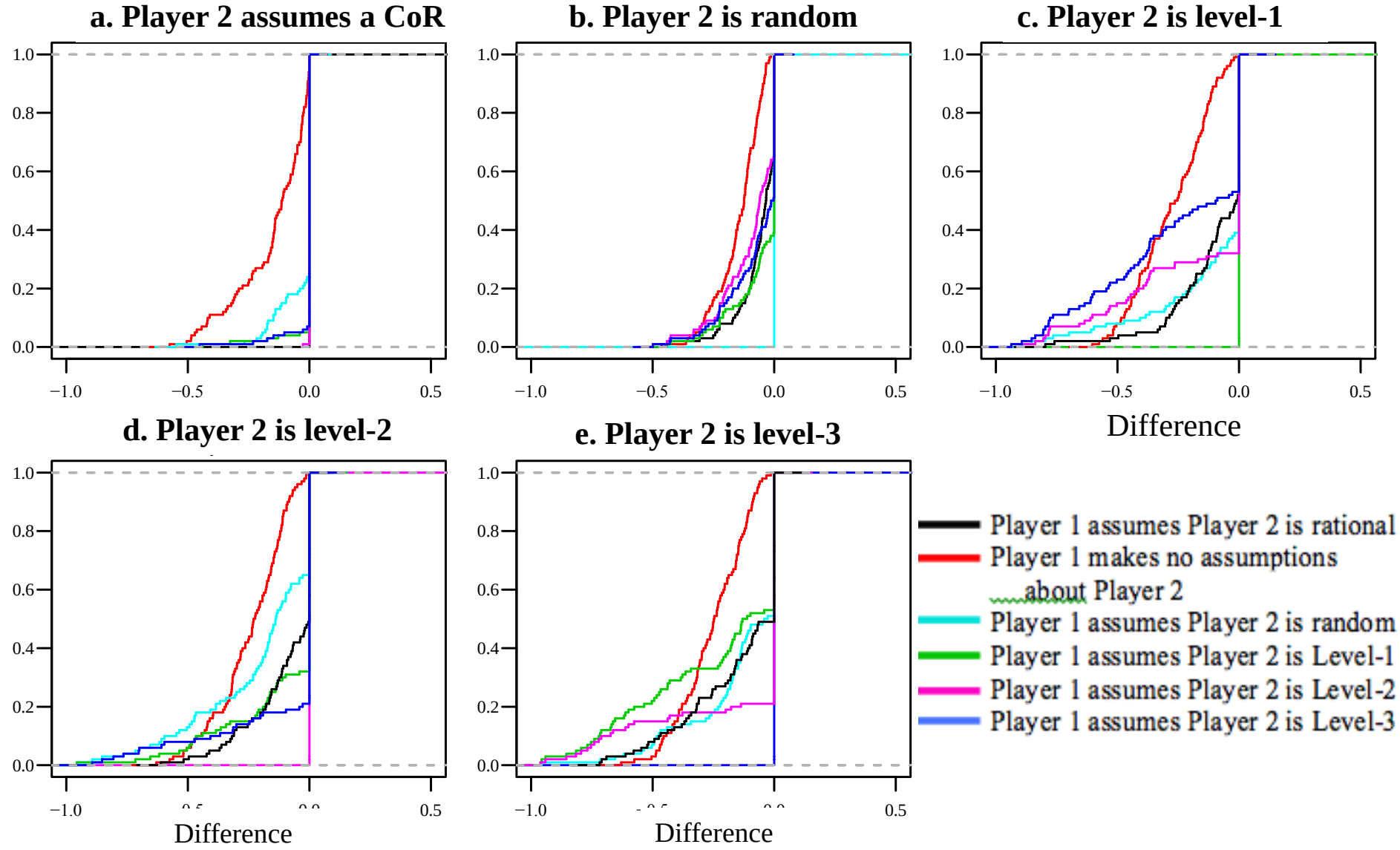


Number of times Player 1's assumption about P2 is <u>best</u> (equivalent to "correct" assumption)

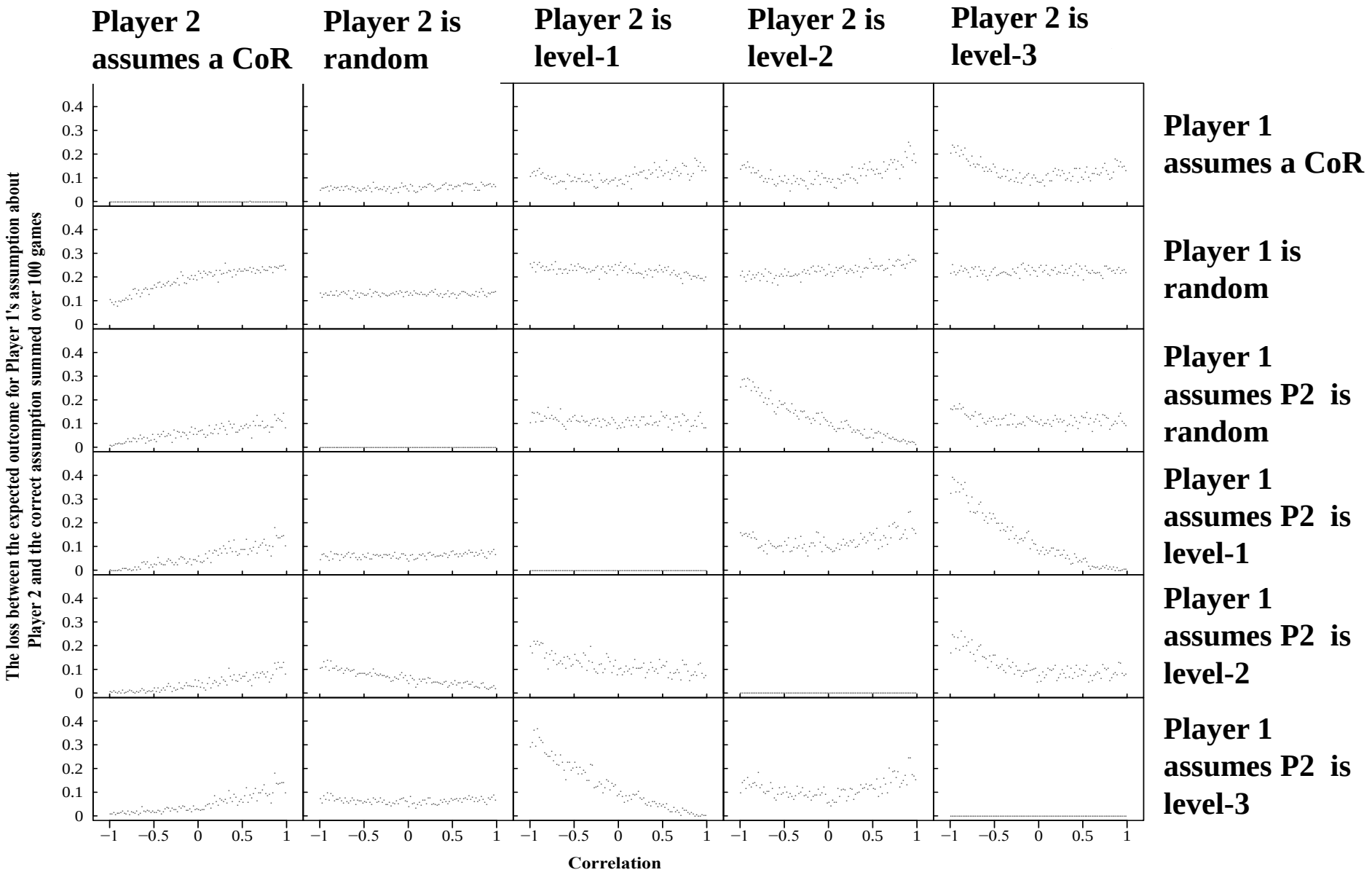


Distribution of differences between the expected payoff and payoff for "correct" assumption

(correlation: -0.70 to -0.72)



Difference between the expected payoff given P1's assumption and the "correct" assumption



Concluding Thoughts

- Assuming a “community of rationality” is not advantageous (and certainly not conservative)
  - Defender may be worse off for making assumption
- Correlations between payoff matrices heavily influences results
- Principles of game theory can be applied to risk assessments, but cautiously (It’s a tool!)
  - Implications of assumptions must be understood
  - Black swans cannot be found
  - Address probabilistic outcomes over the consequence space

Acknowledgements

This work was supported by the National Science Foundation Grants *1149460* and *1331399*. Any opinions, findings, conclusions, or recommendations expressed in this work are those of the author and do not necessarily reflect the views of the National Science Foundation.