Rational Irrationality

Dan Ventura
Computer Science Department
Brigham Young University
ventura@cs.byu.edu

Abstract

We present a game-theoretic account of irrational agent behavior and define conditions under which irrational behavior may be considered *quasi-rational*. To do so, we use a 2-player, zero-sum strategic game, parameterize the reward structure and study how the value of the game changes with this parameter. We argue that for any "underdog" agent, there is a point at which the asymmetry of the game will provoke the agent to act irrationally. This implies that the non-"underdog" player must therefore also act irrationally even though he has no incentive (in the reward structure) for doing so, which implies, in turn, a meta-level game.

Introduction

Game theory offers an often compelling account of agent behavior under various forms of cooperation and conflict. One of the basic tenets of game theory is that agents will behave rationally, that is, that agents will act to maximize their reward in the game. However, it is well-known that humans do not always act in this way. One natural reason for irrational behavior might be that the agent does not understand the game, and is, in fact, acting rationally given their (incorrect) understanding of the game. But, what if an agent fully understands the game and still acts irrationally? An interesting array of security questions can be asked in this context, particularly in cases of asymmetric conflict such as those seen in recent cases of terrorism and military conflicts. As an extreme example, why would someone detonate a bomb strapped to their own body, or fly a plane into a building? Can we offer an explanation for an agent that chooses to play irrationally, knowing that that is what they are choosing? Is there a kind of rationality to this kind of choice? Can we determine the best response to such behavior?

A second, day-to-day scenario that exhibits this kind of irrational behavior might be the purchase of lottery tickets. A tongue-in-cheek acknowledgement of this is the humorous aphorism that explains that "the lottery is a tax on people who can't do math". When people buy lottery tickets, they are (apparently) happy to risk a small amount of money for a chance at a huge reward, even though an expected value analysis will tell them that this is a bad idea (in other words, not rational). Most people, of course, don't know how to or don't want to compute the expected value of the decision to buy a lottery ticket. But even people that do the analysis sometimes will buy lottery tickets. So, the question is, is the lottery really a tax on people that can't do math, or are they just doing a different kind of math?

In this paper we consider a game theoretic setting involving asymmetric reward structure that suggests the latter and as a result suggests that in asymmetric games of conflict it can be *quasi-rational* for an agent to behave irrationally. And, if this is the case, such an agent's opponent must also then behave (or at least be perceived to be capable of behaving) irrationally as well.

The Domination Game

Consider the following payoff matrix for a 2-player, zero-sum strategic game, which we will call Domination:

$$D = \begin{pmatrix} 1 & -1 \\ -1 & r \end{pmatrix}$$

Player A has two actions (a_1, a_2) corresponding to a choice of rows in the matrix and player B has two actions (b_1, b_2) corresponding to a choice of columns. Solving for an equilibrium point yields a value for the game (in terms of r) of $v = \frac{r-1}{r+3}$ which is obtained by player A playing a mixed strategy with probability $p = \frac{r+1}{r+3}$ of playing action a_1 and by player B playing a mixed strategy with probability $q = \frac{r+1}{r+3}$ of playing action b_1 (see Appendix A for details).

This game can be understood intuitively as follows. As the value of r increases, B is less and less likely to choose action b_2 as such a choice could result in a catastrophic payoff. At the same time, A is less and less likely to choose action a_2 as this would likely result in a loss (since the chance of getting payoff r is increasingly remote). And, in the limit, the equilibrium strategy is the pure (joint) strategy that results in the joint action (a_1, b_1) being chosen with probability 1. If both players are rational, their strategies will approach this pure strategy as r approaches ∞.

For this discussion, a useful way to look at the equilibrium value v is as the sum of a reward π and a risk ρ. Figure 1 shows a decomposition of v into π and ρ as a function of r, where $\pi = p(2q-1)$ and $\rho = (1-p)(-q-qr+r)$ (or, alternatively, $\pi = q(2p-1)$ and $\rho = (1-q)(-p-pr+$

r)—see Appendix B for details). The reward component is the result of the opponent playing action 1, while the risk component is the result of the opponent playing action 2. For the joint rational strategy, almost all of the (expected) payoff comes from the reward component, and, as r increases, this contribution continues to increase. In the limit, $\pi = v$, and $\rho = 0$. In other words, in the limit, there is no chance that the opponent will play action 2.

The Case for Irrationality

For the following, without loss of generality, assume $r > 1$, so that $p > 0.5$ and $q > 0.5$. Let $\boldsymbol{p}^* = [p, 1-p]$, $\boldsymbol{q}^* = [q, 1-q]$ with the equilibrium (joint) solution being $(\boldsymbol{p}^*, \boldsymbol{q}^*)$. For strategy $\boldsymbol{s} = [s, 1-s]$, let $\boldsymbol{x}_s^{\boldsymbol{s}}$ be the (row or column) vector of the payoff matrix selected using strategy $\boldsymbol{s}$ and choosing the action associated with probability s.

Definition 1. Given strategies $\boldsymbol{s}$ and $\boldsymbol{t}$, let $\mathrm{E}^{\boldsymbol{s}}[\boldsymbol{t}|s] = \boldsymbol{t} \cdot \boldsymbol{x}_s^{\boldsymbol{s}}$ represent the expectation over the possible actions associated with $\boldsymbol{t}$, given the action associated with probability s for strategy $\boldsymbol{s}$.

Definition 2. For strategies $\boldsymbol{s} = (s, 1-s)$ and $\boldsymbol{t}$, and assuming without loss of generality that $s > 0.5$ we say $\pi_{\boldsymbol{t}}^{\boldsymbol{s}} = s\mathrm{E}^{\boldsymbol{s}}[\boldsymbol{t}|s]$ is the ***t**-reward* against $\boldsymbol{s}$. It is the expectation over the possible choices of $\boldsymbol{t}$, given the choice of the *higher* probability action associated with $\boldsymbol{s}$.

Definition 3. For strategies $\boldsymbol{s} = (s, 1-s)$ and $\boldsymbol{t}$, and assuming without loss of generality that $s > 0.5$, we say $\rho_{\boldsymbol{t}}^{\boldsymbol{s}} = (1-s)\mathrm{E}^{\boldsymbol{s}}[\boldsymbol{t}|1-s]$ is the ***t**-risk* against $\boldsymbol{s}$. It is the expectation over the possible choices of $\boldsymbol{t}$, given the choice of the *lower* probability action associated with $\boldsymbol{s}$.

In particular, for the Domination game we have (see Appendix B for details):

$$\pi_{\boldsymbol{p}^*}^{\boldsymbol{q}^*} = \pi_{\boldsymbol{q}^*}^{\boldsymbol{p}^*} = \frac{r^2-1}{r^2+6r+9}$$
$$\rho_{\boldsymbol{p}^*}^{\boldsymbol{q}^*} = \rho_{\boldsymbol{q}^*}^{\boldsymbol{p}^*} = \frac{2r-2}{r^2+6r+9}$$

Notice that these quantities can be combined in two natural ways, both of which express the value v of the game, just as in Figure 1:

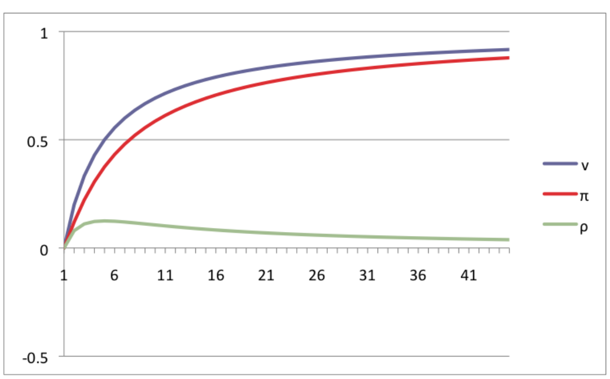

Figure 1: **Decomposing v into reward π and risk ρ, as a function of r**

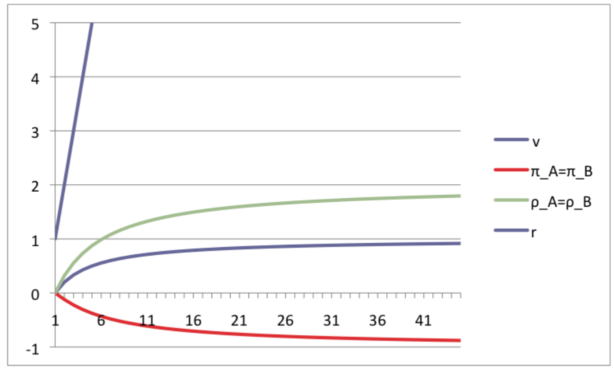

Figure 2: **Irrational reward and risk for A (B) as a function of r**

$$\begin{aligned}
\pi_{\boldsymbol{p}^*}^{\boldsymbol{q}^*} + \rho_{\boldsymbol{p}^*}^{\boldsymbol{q}^*} &= \pi_{\boldsymbol{q}^*}^{\boldsymbol{p}^*} + \rho_{\boldsymbol{q}^*}^{\boldsymbol{p}^*} \\
&= \frac{r^2-1}{r^2+6r+9} + \frac{2r-2}{r^2+6r+9} \\
&= \frac{r^2+2r-3}{r^2+6r+9} \\
&= \frac{r-1}{r+3} \\
&= v
\end{aligned}$$

In other words, v can be understood as the sum of the *equilibrium reward* and *equilibrium risk*. We might alternatively say that v is the sum of the *rational reward* and *rational risk*, and this is the classical solution approach in game theory. However, there are alternative solution approaches that may be considered.

Definition 4. Given a player's equilibrium strategy $\boldsymbol{s}^* = [s, 1-s]$, $\boldsymbol{s}^@ = [1-s, s]$ is that player's *irrational strategy*.

Player A's (B's) irrational strategy for the Domination game is $p^@ = [1-p, p]$ ($q^@ = [1-q, q]$), and when this strategy is played against B's (A's) rational strategy,

$$\pi_{\boldsymbol{p}^@}^{\boldsymbol{q}^*} = \pi_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} = \frac{1-r^2}{r^2+6r+9}$$
$$\rho_{\boldsymbol{p}^@}^{\boldsymbol{q}^*} = \rho_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} = \frac{2r^2+2r-4}{r^2+6r+9}$$

We will sometimes refer to these quantities as *A's (B's) irrational payoff* or *A's (B's) irrational risk*. Figure 2 shows irrational reward and risk for A against B's (and B against A's) rational strategy. Note that the result of playing irrationally is that the reward is now inverted – if the opponent plays rationally, its higher probability action will result in a loss for A, while the risk results in a (possibly very large) win. However, the probability of the risk goes to zero as quickly as the risk grows, resulting in a bounded risk whose expectation is only twice v.

However, since the value $\pi_{\boldsymbol{q}^@}^{\boldsymbol{p}^*}$ is less than v (recall that B prefers smaller payoffs), this is the first scenario we've seen that allows B to actually win (as opposed to just limiting its losses), and it therefore appears to be in B's best interest to act irrationally. How can this be reconciled with the fact that v represents the equilibrium (best for both) value? The answer is because v can now be thought of as also representing an expectation over B's choices given that A chooses rationally (because $\pi^{\boldsymbol{q}^*}_{\boldsymbol{p}^@} = \pi^{\boldsymbol{p}^*}_{\boldsymbol{q}^@}$ and $\rho^{\boldsymbol{q}^*}_{\boldsymbol{p}^@} = \rho^{\boldsymbol{p}^*}_{\boldsymbol{q}^@}$, $\pi^{\boldsymbol{p}^*}_{\boldsymbol{q}^@} + \rho^{\boldsymbol{p}^*}_{\boldsymbol{q}^@} = v$). As r increases, the expectation is skewed more and more by the large value while at the same time, the probability of A actually choosing action a_2 approaches 0. The larger the value of r, the farther away the [irrational] risk ($\rho^{\boldsymbol{p}^*}_{\boldsymbol{q}^@}$) gets from the worst case (r), while the closer the expected [irrational] reward ($\pi^{\boldsymbol{p}^*}_{\boldsymbol{q}^@}$) gets to the best case (-1). So, as r increases, B has (relatively) greater and greater incentive to act irrationally. In other words, the greater A's advantage becomes, the greater incentive B has to act irrationally.

Note that if either one of the players plays irrationally (while the other plays rationally), the value of the game does not change and is still the sum of a reward and a risk,

$$\begin{aligned}
\pi^{\boldsymbol{q}^*}_{\boldsymbol{p}^@} + \rho^{\boldsymbol{q}^*}_{\boldsymbol{p}^@} = \pi^{\boldsymbol{p}^*}_{\boldsymbol{q}^@} + \rho^{\boldsymbol{p}^*}_{\boldsymbol{q}^@} &= \frac{1-r^2}{r^2+6r+9} + \frac{2r^2+2r-4}{r^2+6r+9} \\
&= \frac{r^2+2r-3}{r^2+6r+9} \\
&= \frac{r-1}{r+3} \\
&= v
\end{aligned}$$

As a result, unilateral irrationality does not affect the value v of the game, and these quantities can again be combined to yield two more expressions for v. In other words, v can also be understood as the sum of an *irrational expected payoff* and an *irrational expected risk*, given that one of the players is acting rationally.

To summarize, v can be interpreted in any of three ways: as the expected result for the joint strategy $(\boldsymbol{p}^*, \boldsymbol{q}^*)$ (both players acting rationally), or as the expected result for the joint strategy $(\boldsymbol{p}^*, \boldsymbol{q}^@)$ (player 1 acting rationally and player 2 acting irrationally) or as the expected result for the joint strategy $(\boldsymbol{p}^@, \boldsymbol{q}^*)$ (player 1 acting irrationally and player 2 acting rationally).

The additional joint strategy $(\boldsymbol{p}^@, \boldsymbol{q}^@)$ may also be expanded as above; however, the result cannot be used to express v (see Appendix B for details.)

Definition 5. $\pi^{\#}_{\boldsymbol{s}}$ is the *maximal reward* for strategy $\boldsymbol{s}$. Here *maximal* refers to the amplitude of the payoff, with the sign determined by the goal of the player using $\boldsymbol{s}$.

Definition 6. $\rho^{\#}_{\boldsymbol{s}}$ is the *maximal risk* for strategy $\boldsymbol{s}$. Here *maximal* refers to the amplitude of the payoff, with the sign determined by the goal of the player using $\boldsymbol{s}$.

In a zero-sum game with players using strategies $\boldsymbol{s}$ and $\boldsymbol{t}$, $\rho^{\#}_{\boldsymbol{s}} = \pi^{\#}_{\boldsymbol{t}}$ and $\rho^{\#}_{\boldsymbol{t}} = \pi^{\#}_{\boldsymbol{s}}$, and, in the Domination game, $\pi^{\#}_{\boldsymbol{p}^*} = \pi^{\#}_{\boldsymbol{p}^@} = r$, $\pi^{\#}_{\boldsymbol{q}^*} = \pi^{\#}_{\boldsymbol{q}^@} = -1$, $\rho^{\#}_{\boldsymbol{p}^*} = \rho^{\#}_{\boldsymbol{p}^@} = -1$ and $\rho^{\#}_{\boldsymbol{q}^*} = \rho^{\#}_{\boldsymbol{q}^@} = r$.

Definition 7. $\phi^{\boldsymbol{t}}_{\boldsymbol{s}} = \frac{\pi^{\boldsymbol{t}}_{\boldsymbol{s}}}{\pi^{\#}_{\boldsymbol{s}}}$ is the *reward coefficient* of strategy $\boldsymbol{s}$ against $\boldsymbol{t}$.

Definition 8. $\psi^{\boldsymbol{t}}_{\boldsymbol{s}} = \frac{\rho^{\boldsymbol{t}}_{\boldsymbol{s}}}{\rho^{\#}_{\boldsymbol{s}}}$ is the *risk coefficient* of strategy $\boldsymbol{s}$ against $\boldsymbol{t}$.

For the Domination game, if both players are playing rationally (joint strategy $(\boldsymbol{p}^*, \boldsymbol{q}^*)$),

$$\begin{aligned}
\phi^{\boldsymbol{q}^*}_{\boldsymbol{p}^*} &= \frac{\pi^{\boldsymbol{q}^*}_{\boldsymbol{p}^*}}{\pi^{\#}_{\boldsymbol{p}^*}} = \frac{\frac{r^2-1}{r^2+6r+9}}{r} = \frac{r^2-1}{r^3+6r^2+9r} \\
\psi^{\boldsymbol{q}^*}_{\boldsymbol{p}^*} &= \frac{\rho^{\boldsymbol{q}^*}_{\boldsymbol{p}^*}}{\rho^{\#}_{\boldsymbol{p}^*}} = \frac{\frac{2r-2}{r^2+6r+9}}{-1} = \frac{2-2r}{r^2+4r+4} \\
\phi^{\boldsymbol{p}^*}_{\boldsymbol{q}^*} &= \frac{\pi^{\boldsymbol{p}^*}_{\boldsymbol{q}^*}}{\pi^{\#}_{\boldsymbol{q}^*}} = \frac{\frac{r^2-1}{r^2+6r+9}}{-1} = \frac{1-r^2}{r^2+6r+9} \\
\psi^{\boldsymbol{p}^*}_{\boldsymbol{q}^*} &= \frac{\rho^{\boldsymbol{p}^*}_{\boldsymbol{q}^*}}{\rho^{\#}_{\boldsymbol{q}^*}} = \frac{\frac{2r-2}{r^2+6r+9}}{r} = \frac{2r-2}{r^3+6r^2+9r}
\end{aligned}$$

For the Domination game, if player A is playing irrationally while player B is playing rationally (joint strategy $(\boldsymbol{p}^@, \boldsymbol{q}^*)$),

$$\begin{aligned}
\phi^{\boldsymbol{q}^*}_{\boldsymbol{p}^@} &= \frac{\pi^{\boldsymbol{q}^*}_{\boldsymbol{p}^@}}{\pi^{\#}_{\boldsymbol{p}^@}} = \frac{\frac{1-r^2}{r^2+6r+9}}{r} = \frac{1-r^2}{r^3+6r^2+9r} \\
\psi^{\boldsymbol{q}^*}_{\boldsymbol{p}^@} &= \frac{\rho^{\boldsymbol{q}^*}_{\boldsymbol{p}^@}}{\rho^{\#}_{\boldsymbol{p}^@}} = \frac{\frac{2r^2+2r-4}{r^2+6r+9}}{-1} = \frac{-2r^2-2r+4}{r^2+4r+4}
\end{aligned}$$

On the other hand, if player A is playing rationally while player B is playing irrationally (joint strategy $(\boldsymbol{p}^*, \boldsymbol{q}^@)$),

$$\begin{aligned}
\phi^{\boldsymbol{p}^*}_{\boldsymbol{q}^@} &= \frac{\pi^{\boldsymbol{p}^*}_{\boldsymbol{q}^@}}{\pi^{\#}_{\boldsymbol{q}^@}} = \frac{\frac{1-r^2}{r^2+6r+9}}{-1} = \frac{r^2-1}{r^2+6r+9} \\
\psi^{\boldsymbol{p}^*}_{\boldsymbol{q}^@} &= \frac{\rho^{\boldsymbol{p}^*}_{\boldsymbol{q}^@}}{\rho^{\#}_{\boldsymbol{q}^@}} = \frac{\frac{2r^2+2r-4}{r^2+6r+9}}{r} = \frac{2r^2+2r-4}{r^3+6r^2+9r}
\end{aligned}$$

Definition 9. Given a rational (equilibrium) joint strategy $(\boldsymbol{s}^*, \boldsymbol{t}^*)$ for players A and B, where $\boldsymbol{s}^* = [s, 1-s]$, say it is *quasi-rational* for A to play the irrational strategy $\boldsymbol{s}^@ = [1-s, s]$ against $\boldsymbol{t}^*$ if the following conditions hold:

1. $\pi^{\boldsymbol{t}^*}_{\boldsymbol{s}^@} > \pi^{\boldsymbol{t}^*}_{\boldsymbol{s}^*}$ (for player B, who wants to minimize payoff, the inequality is reversed)
2. $\phi^{\boldsymbol{t}^*}_{\boldsymbol{s}^@} > \theta^{\phi}_{A}$
3. $\psi^{\boldsymbol{t}^*}_{\boldsymbol{s}^@} < \theta^{\psi}_{A}$

The first condition ensures that the irrational reward is better than the rational reward. The second condition ensures that the irrational reward is sufficiently close to the maximal reward, where "sufficiently close" is defined by θ^{ϕ}_{A}. The third condition ensures that the irrational risk is significantly less than the maximal risk, where "significantly less" is defined by θ^{ψ}_{A}. We will refer to θ^{ϕ}_{A} as A's *reward threshold* and to θ^{ψ}_{A} as A's *risk threshold*. A reckless player will have a low reward threshold and a high risk threshold. A cautious player will be opposite. Both thresholds low would characterize a nickel-and-dime player, while both high would characterize a player that pursues high-risk/high-reward scenarios. Note that because $\forall_{s,t} -1 \leq \pi_s^t, \rho_s^t \leq 1$, it is sufficient that $-1 \leq \theta_A^\phi, \theta_A^\psi \leq 1$.

Definition 10. A *rational game* is one for which the quasi-rationality conditions do not hold for either player.

Definition 11. A *quasi-rational game* is one for which the quasi-rationality conditions hold for exactly one player.

Definition 12. An *irrational game* is one for which the quasi-rationality conditions hold for both players.

Proposition 1. *Lots of games are rational, including any whose reward structures are not asymmetric.*

Theorem 1. *For any two players A and B, with thresholds $0 < \theta_A^\phi, \theta_A^\psi, \theta_B^\phi, \theta_B^\psi < 1$, there is a game G that is quasi-rational.*

Proof. Let G be the Domination game with $r > \max\{\frac{-3\theta_B^\phi - \sqrt{8\theta_B^\phi+1}}{\theta_B^\phi - 1}, \frac{2}{\theta_B^\psi}\}$. We must show that the conditions of quasi-rationality hold for exactly one player. Without loss of generality, let that player be B. By Lemma 1, we have that the quasi-rationality conditions hold for B, and by Lemma 2, we have that they do not hold for A. □

Lemma 1. *Let $(\boldsymbol{p}^*, \boldsymbol{q}^*)$ be the equilibrium strategy for A and B for the Domination game. Given thresholds $0 < \theta_B^\phi, \theta_B^\psi < 1$, it is quasi-rational for B to play $\boldsymbol{q}^@$ against A for $r > \max\{\frac{-3\theta_B^\phi - \sqrt{8\theta_B^\phi+1}}{\theta_B^\phi - 1}, \frac{2}{\theta_B^\psi}\}$.*

Proof. For the first condition we must show $\pi_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} < \pi_{\boldsymbol{q}^*}^{\boldsymbol{p}^*}$. (note that since B wants to minimize payoff, we have reversed the inequality.

$$\begin{aligned}
\pi_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} &= \frac{1-r^2}{r^2+6r+9} \\
&< \frac{r^2-1}{r^2+6r+9} \\
&= \pi_{\boldsymbol{q}^*}^{\boldsymbol{p}^*}
\end{aligned}$$

Note that the inequality holds because $\theta_B^\phi > 0$ implies $r^2 > 1$.

For the second condition we must show $\phi_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} > \theta_B^\phi$.

$$\begin{aligned}
\phi_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} &= \frac{r^2-1}{r^2+6r+9} \\
&> \frac{\left(\frac{-3\theta_B^\phi - \sqrt{8\theta_B^\phi+1}}{\theta_B^\phi-1}\right)^2 - 1}{\left(\frac{-3\theta_B^\phi - \sqrt{8\theta_B^\phi+1}}{\theta_B^\phi-1}\right)^2 + 6\left(\frac{-3\theta_B^\phi - \sqrt{8\theta_B^\phi+1}}{\theta_B^\phi-1}\right) + 9} \\
&= \frac{\frac{9(\theta_B^\phi)^2 + 6\theta_B^\phi\sqrt{8\theta_B^\phi+1} + 8\theta_B^\phi + 1}{(\theta_B^\phi-1)^2} - 1}{\frac{9(\theta_B^\phi)^2 + 6\theta_B^\phi\sqrt{8\theta_B^\phi+1} + 8\theta_B^\phi + 1}{(\theta_B^\phi-1)^2} + \frac{-18\theta_B^\phi - 6\sqrt{8\theta_B^\phi+1}}{(\theta_B^\phi-1)} + 9} \\
&= \frac{8(\theta_B^\phi)^2 + 6\theta_B^\phi\sqrt{8\theta_B^\phi+1} + 10\theta_B^\phi}{8\theta_B^\phi + 6\sqrt{8\theta_B^\phi+1} + 10} \\
&= \frac{\theta_B^\phi(8\theta_B^\phi + 6\sqrt{8\theta_B^\phi+1} + 10)}{8\theta_B^\phi + 6\sqrt{8\theta_B^\phi+1} + 10} \\
&= \theta_B^\phi
\end{aligned}$$

For the third condition we must show $\psi_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} < \theta_B^\psi$.

$$\begin{aligned}
\psi_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} &= \frac{2r^2+2r-4}{r^3+6r^2+9r} \\
&< \frac{2\left(\frac{2}{\theta_B^\psi}\right)^2 + 2\left(\frac{2}{\theta_B^\psi}\right) - 4}{\left(\frac{2}{\theta_B^\psi}\right)^3 + 6\left(\frac{2}{\theta_B^\psi}\right)^2 + 9\left(\frac{2}{\theta_B^\psi}\right)} \\
&= \frac{2\left(\frac{\theta_B^\psi}{2}\right) + 2\left(\frac{\theta_B^\psi}{2}\right)^2 - 4\left(\frac{\theta_B^\psi}{2}\right)^3}{1 + 6\left(\frac{\theta_B^\psi}{2}\right) + 9\left(\frac{\theta_B^\psi}{2}\right)^2} \\
&= \frac{\frac{\theta_B^\psi}{2}\left(2-\theta_B^\psi\right)\left(1+\theta_B^\psi\right)}{\left(1+\frac{3}{2}\theta_B^\psi\right)\left(1+\frac{3}{2}\theta_B^\psi\right)} \\
&= \frac{\theta_B^\psi}{2}\frac{\left(2-\theta_B^\psi\right)}{\left(1+\frac{3}{2}\theta_B^\psi\right)}\frac{\left(1+\theta_B^\psi\right)}{\left(1+\frac{3}{2}\theta_B^\psi\right)} \\
&< \frac{\theta_B^\psi}{2} \cdot 2 \cdot 1 \\
&= \theta_B^\psi
\end{aligned}$$

□

Lemma 2. *Let $(\boldsymbol{p}^*, \boldsymbol{q}^*)$ be the equilibrium strategy for A and B for the Domination game. It is not quasi-rational for A to play $\boldsymbol{p}^@$ against B for $r > 1$.*

Proof. Since A wants to maximize payoff, it suffices to show that $\pi_{\boldsymbol{p}^{@}}^{\boldsymbol{q}^*} \leq \pi_{\boldsymbol{p}^*}^{\boldsymbol{q}^*}$.

$$\begin{aligned} \pi_{\boldsymbol{p}^{@}}^{\boldsymbol{q}^*} &= \frac{1-r^2}{r^2+6r+9} \\ &< \frac{r^2-1}{r^2+6r+9} \\ &= \pi_{\boldsymbol{p}^*}^{\boldsymbol{q}^*} \end{aligned}$$

□

Proposition 2. *No irrational one-shot, zero-sum strategic game exists.*

The Meta-game Insanity

Theorem 1 says that for even the most conservative "underdog" and the most reckless "bully", there exists a game for which the "underdog" should play irrationally, but the "bully" should not. Of course, if the "underdog" *does* play irrationally, then the "bully" would benefit by playing irrationally, and if the "bully" chooses to play irrationally then the "underdog" could not afford to. This, in turn, suggests a meta-game, which we will call *Insanity*, in which the choice of row and column determine the players' choices of whether to play rationally or irrationally for the Domination game. The payoff matrix can be represented as the (expected) value of the game given both player's choice of strategy:

$$I = \begin{pmatrix} v & -v \\ -v & rv \end{pmatrix} = vD$$

The upper left represents the expected value of the game for both players acting rationally and the bottom right for both acting irrationally. The bottom left is player A acting irrationally and B acting rationally and the upper right is A acting rationally and B acting irrationally. Note that the form of the matrix is exactly the same as that of the Domination game, and, indeed, the equilibrium strategies, $\tilde{\boldsymbol{p}}^* = [\tilde{p}, 1-\tilde{p}]$ and $\tilde{\boldsymbol{q}}^* = [\tilde{q}, 1-\tilde{q}]$ are the same as those for the Domination game (see Appendix C): $\tilde{p} = p$, $\tilde{q} = q$, and given these, we can compute the equilibrium value of the Insanity game (Appendix C): $\tilde{v} = v^2$, which is obtained by player A playing (meta-)strategy $\tilde{\boldsymbol{p}}^*$ and player B playing (meta-)strategy $\tilde{\boldsymbol{q}}^*$. What this means is that player A plays a mixed (meta-)strategy with probability $\tilde{p}$ of playing rationally (that is, employing strategy $\boldsymbol{p}^*$) for the Domination game and probability $1-\tilde{p}$ of playing irrationally (strategy $\boldsymbol{p}^{@}$ for the Domination game) while player B plays a similar mixed (meta-)strategy over the strategies $\boldsymbol{q}^*$ and $\boldsymbol{q}^{@}$.

We can now ask, given the meta-strategies, how this affects the probabilities of the players choosing between their two (base) action choices. If p is the probability that A will choose action a_1 in the base game, we will call p_1 the probability that A will choose a_1 after first playing the meta-game (and thus selecting whether or not to play the base game rationally). If A chooses to play rationally (probability $\tilde{p}$), then A will play a_1 with probability p. If A chooses to play irrationally (probability $1-\tilde{p}$), then A will play a_1 with probability $1-p$. So, $p_1 = \tilde{p}p + (1-\tilde{p})(1-p)$. Likewise, q_1 will represent the probability that B will choose action b_1, and $q_1 = \tilde{q}q + (1-\tilde{q})(1-q)$. With these, we can derive an alternative form of the value of the meta-game Insanity, that we will call v_1 (see Appendix D for details):

$$\begin{aligned} v_1 &= p_1q_1(1) + (1-p_1)q_1(-1) \\ &\quad + p_1(1-q_1)(-1) \\ &\quad +(1-p_1)(1-q_1)(r) \\ &= v^2 \end{aligned}$$

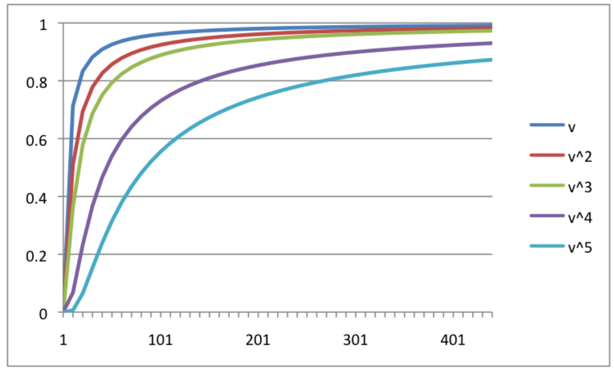

Figure 3: **Value of the Domination game as a function of r for various levels of meta-gamesmanship**

It should be obvious that this meta-gamesmanship can be carried on *ad infinitum*, and we can identify the meta-level with the subscript k (we have been discussing the first meta-level with $k = 1$). A little thought should verify that the value of the kth level game is

$$v_k = v^{k+1}$$

Figure 3 shows how the value of the game changes for various levels of meta-gamesmanship. While the ultimate value of the game is always the same in the limit (1), the threat of irrationality can have a significant effect on the value for $r << \infty$.

Further, the probability that A will eventually play a_1 given his choice of strategy at the kth level can be defined recursively as

$$\begin{aligned} p_0 &= p \\ p_k &= pp_{k-1} + (1-p)(1-p_{k-1}), k > 0 \end{aligned}$$

with q_k defined similarly. Figure 4 shows how the probability p of A playing a_1 changes for various levels of metagamemanship. As expected, as k increases, p grows more slowly. Again, while the asymptotic results remain unchanged, the threat of irrationality by B affects the behavior of player A.

Final Comments

The ideas suggested here are underdeveloped and much work remains to be done. Most obviously, experiments must be designed to place people in situations similar to the Domination game to see if, in fact, they do behave as predicted (that is, they behave quasi-rationally when the game

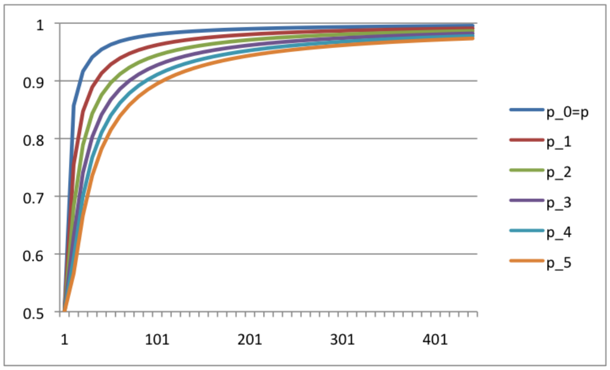

Figure 4: **Probability p of A playing a_1 as a function of r for various levels of meta-gamesmanship**

is significantly asymmetric). We might also be able to make some empirical estimation of general values for reward and risk thresholds. In addition, other work on irrationality in game theory has been proposed, though much of it appears to treat the case of extensive games (Krepps et al. 1982; Neyman 1985; Pettit and Sugden 1989; Aumann 1992). Can any of this be applied to the strategic case? Or, can (and should) this work be extended to extensive games so that a comparison with these theories can be attempted?

References

Aumann, R. 1992. Irrationality in game theory. *Economic Analysis of Markets and Games, Essays in Honor of Frank Hahn* 214–227.

Krepps, D.; Milgrom, P.; Roberts, J.; and Wilson, R. 1982. Rational cooperation in the finitely repeated prisoners dilemma. *Journal of Economic Theory* 27:245–252.

Neyman, A. 1985. Bounded complexity justifies cooperation in the finitely repeated prisoner's dilemma. *Economics Letters* 19:227–229.

Pettit, P., and Sugden, R. 1989. The backward induction paradox. *The Journal of Philosophy* 86:169–182.

Appendices

A Equilibrium of Domination game

To find A's equilibrium strategy, we compute

$$\begin{aligned} \mathrm{E}[p|q=1] &= p(1)+(1-p)(-1) = 2p-1 \\ \mathrm{E}[p|q=0] &= p(-1)+(1-p)(r) = -p-pr+r \end{aligned}$$

Settting these equal and solving for p gives A's equilibrium strategy $\boldsymbol{p}^* = (p, 1-p)$:

$$\begin{aligned} 2p-1 &= -p-pr+r \\ 3p+pr &= r+1 \\ p(r+3) &= r+1 \\ p &= \frac{r+1}{r+3} \end{aligned}$$

To find B's equilibrium strategy, we compute

$$\begin{aligned} \mathrm{E}[q|p=1] &= q(1)+(1-q)(-1) = 2q-1 \\ \mathrm{E}[q|p=0] &= q(-1)+(1-q)(r) = -q-qr+r \end{aligned}$$

and repeat the process to find $\boldsymbol{q}^* = (q, 1-q)$:

$$\begin{aligned} 2q-1 &= -q-qr+r \\ 3q+qr &= r+1 \\ q(r+3) &= r+1 \\ q &= \frac{r+1}{r+3} \end{aligned}$$

The (equilibrium) value v of the game can then be expressed as the expected payoff, $\mathrm{E}[(p,1-p),(q,1-q)]$, when both players play their equilibrium strategies:

$$\begin{aligned} v &= \mathrm{E}[(p,1-p),(q,1-q)] \\ &= pq(1)+(1-p)q(-1)+p(1-q)(-1) \\ &\quad +(1-p)(1-q)(r) \\ &= pq-(1-p)q-p(1-q)+(1-p)(1-q)r \\ &= \frac{r^2+2r+1}{r^2+6r+9} - \frac{2r+2}{r^2+6r+9} - \frac{2r+2}{r^2+6r+9} \\ &\quad + \frac{4}{r^2+6r+9}r \\ &= \frac{r^2+2r-3}{r^2+6r+9} \\ &= \frac{r-1}{r+3} \end{aligned}$$

B Decomposing value of Domination game into risk and reward

For convenience, we note the following equalities, which are easily verified. The first group is expectations based on bipartisan rationality:

$$\begin{aligned} \mathrm{E}^{\boldsymbol{q}^*}[\boldsymbol{p}^*|q] &= \boldsymbol{p}^* \cdot \boldsymbol{x}_q^{\boldsymbol{q}^*} = [p,(1-p)][1,-1]^T \\ &= 2p-1 = \frac{r-1}{r+3} \\ \mathrm{E}^{\boldsymbol{q}^*}[\boldsymbol{p}^*|1-q] &= \boldsymbol{p}^* \cdot \boldsymbol{x}_{1-q}^{\boldsymbol{q}^*} = [p,(1-p)][-1,r]^T \\ &= -p+(1-p)r = \frac{r-1}{r+3} \\ \mathrm{E}^{\boldsymbol{p}^*}[\boldsymbol{q}^*|p] &= \boldsymbol{q}^* \cdot \boldsymbol{x}_p^{\boldsymbol{p}^*} = [q,(1-q)][1,-1]^T \\ &= 2q-1 = \frac{r-1}{r+3} \\ \mathrm{E}^{\boldsymbol{p}^*}[\boldsymbol{q}^*|1-p] &= \boldsymbol{q}^* \cdot \boldsymbol{x}_{1-p}^{\boldsymbol{p}^*} = [q,(1-q)][-1,r]^T \\ &= -q+(1-q)r = \frac{r-1}{r+3} \end{aligned}$$

The second group is expectations based on unilateral irrationality:

$$\begin{aligned}
\mathrm{E}^{\boldsymbol{q}^*}[\boldsymbol{p}^@|q] &= \boldsymbol{p}^@ \cdot \boldsymbol{x}_q^{\boldsymbol{q}^*} = [(1-p), p][1, -1]^T \\
&= 1 - 2p = \frac{1-r}{r+3} \\
\mathrm{E}^{\boldsymbol{q}^*}[\boldsymbol{p}^@|1-q] &= \boldsymbol{p}^@ \cdot \boldsymbol{x}_{1-q}^{\boldsymbol{q}^*} = [(1-p), p][-1, r]^T \\
&= p + pr - 1 = \frac{r^2+r-2}{r+3} \\
\mathrm{E}^{\boldsymbol{p}^*}[\boldsymbol{q}^@|p] &= \boldsymbol{q}^@ \cdot \boldsymbol{x}_p^{\boldsymbol{p}^*} = [(1-q), q][1, -1]^T \\
&= 1 - 2q = \frac{1-r}{r+3} \\
\mathrm{E}^{\boldsymbol{p}^*}[\boldsymbol{q}^@|1-p] &= \boldsymbol{q}^@ \cdot \boldsymbol{x}_{1-p}^{\boldsymbol{p}^*} = [(1-q), q][-1, r]^T \\
&= q + qr - 1 = \frac{r^2+r-2}{r+3}
\end{aligned}$$

And the last group is expectations based on bipartisan irrationality:

$$\begin{aligned}
\mathrm{E}^{\boldsymbol{q}^@}[\boldsymbol{p}^@|q] &= \boldsymbol{p}^@ \cdot \boldsymbol{x}_q^{\boldsymbol{q}^@} = [(1-p), p][-1, r]^T \\
&= p + pr - 1 = \frac{r^2+r-2}{r+3} \\
\mathrm{E}^{\boldsymbol{q}^@}[\boldsymbol{p}^@|1-q] &= \boldsymbol{p}^@ \cdot \boldsymbol{x}_{1-q}^{\boldsymbol{q}^@} = [(1-p), p][1, -1]^T \\
&= 1 - 2p = \frac{1-r}{r+3} \\
\mathrm{E}^{\boldsymbol{p}^@}[\boldsymbol{q}^@|p] &= \boldsymbol{q}^@ \cdot \boldsymbol{x}_p^{\boldsymbol{p}^@} = [(1-q), q][-1, r]^T \\
&= q + qr - 1 = \frac{r^2+r-2}{r+3} \\
\mathrm{E}^{\boldsymbol{p}^@}[\boldsymbol{q}^@|1-p] &= \boldsymbol{q}^@ \cdot \boldsymbol{x}_{1-p}^{\boldsymbol{p}^@} = [(1-q), q][1, -1]^T \\
&= 1 - 2q = \frac{1-r}{r+3}
\end{aligned}$$

Given the first group, we can now compute rewards and risks for bipartisan rationality:

$$\begin{aligned}
\pi_{\boldsymbol{p}^*}^{\boldsymbol{q}^*} &= q\mathrm{E}^{\boldsymbol{q}^*}[\boldsymbol{p}^*|q] = \left(\frac{r+1}{r+3}\right)\left(\frac{r-1}{r+3}\right) = \frac{r^2-1}{r^2+6r+9} \\
\rho_{\boldsymbol{p}^*}^{\boldsymbol{q}^*} &= (1-q)\mathrm{E}^{\boldsymbol{q}^*}[\boldsymbol{p}^*|1-q] = \left(\frac{2}{r+3}\right)\left(\frac{r-1}{r+3}\right) \\
&= \frac{2r-2}{r^2+6r+9} \\
\pi_{\boldsymbol{q}^*}^{\boldsymbol{p}^*} &= p\mathrm{E}^{\boldsymbol{p}^*}[\boldsymbol{q}^*|p] = \left(\frac{r+1}{r+3}\right)\left(\frac{r-1}{r+3}\right) = \frac{r^2-1}{r^2+6r+9} \\
\rho_{\boldsymbol{q}^*}^{\boldsymbol{p}^*} &= (1-p)\mathrm{E}^{\boldsymbol{p}^*}[\boldsymbol{q}^*|1-p] = \left(\frac{2}{r+3}\right)\left(\frac{r-1}{r+3}\right) \\
&= \frac{2r-2}{r^2+6r+9}
\end{aligned}$$

and with the second, for unilateral irrationality:

$$\begin{aligned}
\pi_{\boldsymbol{p}^@}^{\boldsymbol{q}^*} &= q\mathrm{E}^{\boldsymbol{q}^*}[\boldsymbol{p}^@|q] = \left(\frac{r+1}{r+3}\right)\left(\frac{1-r}{r+3}\right) = \frac{1-r^2}{r^2+6r+9} \\
\rho_{\boldsymbol{p}^@}^{\boldsymbol{q}^*} &= (1-q)\mathrm{E}^{\boldsymbol{q}^*}[\boldsymbol{p}^@|1-q] = \left(\frac{2}{r+3}\right)\left(\frac{r^2+r-2}{r+3}\right) \\
&= \frac{2r^2+2r-4}{r^2+6r+9} \\
\pi_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} &= p\mathrm{E}^{\boldsymbol{p}^*}[\boldsymbol{q}^@|p] = \left(\frac{r+1}{r+3}\right)\left(\frac{1-r}{r+3}\right) = \frac{1-r^2}{r^2+6r+9} \\
\rho_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} &= (1-p)\mathrm{E}^{\boldsymbol{p}^*}[\boldsymbol{q}^@|1-p] = \left(\frac{2}{r+3}\right)\left(\frac{r^2+r-2}{r+3}\right) \\
&= \frac{2r^2+2r-4}{r^2+6r+9}
\end{aligned}$$

and with the third, the case for bipartisan irrationality:

$$\begin{aligned}
\pi_{\boldsymbol{p}^@}^{\boldsymbol{q}^@} &= q\mathrm{E}^{\boldsymbol{q}^@}[\boldsymbol{p}^@|q] = \left(\frac{r+1}{r+3}\right)\left(\frac{r^2+r-2}{r+3}\right) \\
&= \frac{r^3+2r^2-r-2}{r^2+6r+9} \\
\rho_{\boldsymbol{p}^@}^{\boldsymbol{q}^@} &= (1-q)\mathrm{E}^{\boldsymbol{q}^@}[\boldsymbol{p}^@|1-q] = \left(\frac{2}{r+3}\right)\left(\frac{1-r}{r+3}\right) \\
&= \frac{2-2r}{r^2+6r+9} \\
\pi_{\boldsymbol{q}^@}^{\boldsymbol{p}^@} &= p\mathrm{E}^{\boldsymbol{p}^@}[\boldsymbol{q}^@|p] = \left(\frac{r+1}{r+3}\right)\left(\frac{r^2+r-2}{r+3}\right) \\
&= \frac{r^3+2r^2-r-2}{r^2+6r+9} \\
\rho_{\boldsymbol{q}^@}^{\boldsymbol{p}^@} &= (1-p)\mathrm{E}^{\boldsymbol{p}^@}[\boldsymbol{q}^@|1-p] = \left(\frac{2}{r+3}\right)\left(\frac{1-r}{r+3}\right) \\
&= \frac{2-2r}{r^2+6r+9}
\end{aligned}$$

Finally, we can see that for bipartisan rationality, the value v of the game is the sum of the rational reward and the rational risk (from either player's perspective):

$$\begin{aligned}
\pi_{\boldsymbol{p}^*}^{\boldsymbol{q}^*} + \rho_{\boldsymbol{p}^*}^{\boldsymbol{q}^*} = \pi_{\boldsymbol{q}^*}^{\boldsymbol{p}^*} + \rho_{\boldsymbol{q}^*}^{\boldsymbol{p}^*} &= \frac{r^2-1}{r^2+6r+9} + \frac{2r-2}{r^2+6r+9} \\
&= \frac{r^2+2r-3}{r^2+6r+9} \\
&= \frac{r-1}{r+3} \\
&= v
\end{aligned}$$

and that, in fact, the same thing may be said for either case of unilateral irrationality:

$$\begin{aligned}
\pi_{\boldsymbol{p}^@}^{\boldsymbol{q}^*} + \rho_{\boldsymbol{p}^@}^{\boldsymbol{q}^*} = \pi_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} + \rho_{\boldsymbol{q}^@}^{\boldsymbol{p}^*} &= \frac{1-r^2}{r^2+6r+9} + \frac{2r^2+2r-4}{r^2+6r+9} \\
&= \frac{r^2+2r-3}{r^2+6r+9} \\
&= \frac{r-1}{r+3} \\
&= v
\end{aligned}$$

However, bipartisan irrationality does not conserve the game's value, instead scaling it by r:

$$\begin{aligned}
\pi_{\boldsymbol{p}^{@}}^{\boldsymbol{q}^{@}} + \rho_{\boldsymbol{p}^{@}}^{\boldsymbol{q}^{@}} &= \pi_{\boldsymbol{q}^{@}}^{\boldsymbol{p}^{@}} + \rho_{\boldsymbol{q}^{@}}^{\boldsymbol{p}^{@}} \\
&= \frac{r^3 + 2r^2 - r - 2}{r^2 + 6r + 9} + \frac{2 - 2r}{r^2 + 6r + 9} \\
&= \frac{r^3 + 2r^2 - 3r}{r^2 + 6r + 9} \\
&= \frac{r^2 - r}{r + 3} \\
&= \frac{r(r - 1)}{r + 3} \\
&= rv
\end{aligned}$$

C Equilibrium of Insanity game

To find A's equilibrium strategy, we compute

$$\mathrm{E}[\tilde{p}|\tilde{q} = 1] = \tilde{p}v - (1 - \tilde{p})v$$

$$\mathrm{E}[\tilde{p}|\tilde{q} = 0] = -\tilde{p}v + (1 - \tilde{p})rv$$

Setting these equal and solving for $\tilde{p}$ gives A's equilibrium strategy $\boldsymbol{p}^* = (\tilde{p}, 1 - \tilde{p})$:

$$\begin{aligned}
\tilde{p}v - (1 - \tilde{p})v &= -\tilde{p}v + (1 - \tilde{p})rv \\
\tilde{p}v - v + \tilde{p}v &= -\tilde{p}v + rv - \tilde{p}rv \\
3\tilde{p}v + \tilde{p}rv &= v + rv \\
\tilde{p}(rv + 3v) &= rv + v \\
\tilde{p} &= \frac{rv + v}{rv + 3v} \\
\tilde{p} &= \frac{r + 1}{r + 3}
\end{aligned}$$

Since the game is symmetric, we have again that $\tilde{q} = \tilde{p}$ and B's equilibrium strategy $\boldsymbol{q}_1^* = (\tilde{q}, 1 - \tilde{q})$ is again the same as A's: $\tilde{q} = \frac{r+1}{r+3} = q$. The (equilibrium) value $\tilde{v}$ of the game can then be expressed as the expected payoff, $\mathrm{E}[(\tilde{p}, 1 - \tilde{p}), (\tilde{q}, 1 - \tilde{q})]$, when both players play their equilibrium strategies:

$$\begin{aligned}
\tilde{v} &= \mathrm{E}[(\tilde{p}, 1 - \tilde{p}), (\tilde{q}, 1 - \tilde{q})] \\
&= \tilde{p}\tilde{q}(v) + (1 - \tilde{p})\tilde{q}(-v) + \tilde{p}(1 - \tilde{q})(-v) \\
&\quad + (1 - \tilde{p})(1 - \tilde{q})(rv) \\
&= v(\tilde{p}\tilde{q} - (1 - \tilde{p})\tilde{q} - \tilde{p}(1 - \tilde{q}) + (1 - \tilde{p})(1 - \tilde{q})r) \\
&= v^2
\end{aligned}$$

with the last step possible because $\tilde{p} = p$ and $\tilde{q} = q$ and using the third equality from the derivation of v in Appendix A.

D Revisiting value of Insanity game

$$\begin{aligned}
v_1 &= p_1q_1(1) + (1 - p_1)q_1(-1) + p_1(1 - q_1)(-1) \\
&\quad + (1 - p_1)(1 - q_1)(r) \\
&= [(\tilde{p}p + (1 - \tilde{p})(1 - p))(\tilde{q}q + (1 - \tilde{q})(1 - q))](1) \\
&\quad + [(1 - (\tilde{p}p + (1 - \tilde{p})(1 - p))) \\
&\quad (\tilde{q}q + (1 - \tilde{q})(1 - q))](-1) \\
&\quad + [(\tilde{p}p + (1 - \tilde{p})(1 - p)) \\
&\quad (1 - (\tilde{q}q + (1 - \tilde{q})(1 - q)))](-1) \\
&\quad + [(1 - (\tilde{p}p + (1 - \tilde{p})(1 - p))) \\
&\quad (1 - (\tilde{q}q + (1 - \tilde{q})(1 - q)))](r) \\
&= [(\tilde{p}p + (1 - \tilde{p})(1 - p))(\tilde{q}q + (1 - \tilde{q})(1 - q))](1) \\
&\quad +[(\tilde{p}(1 - p) + (1 - \tilde{p})p)(\tilde{q}q + (1 - \tilde{q})(1 - q))](-1) \\
&\quad +[(\tilde{p}p + (1 - \tilde{p})(1 - p))(\tilde{q}(1 - q) + (1 - \tilde{q})q)](-1) \\
&\quad +[(\tilde{p}(1 - p) + (1 - \tilde{p})p)(\tilde{q}(1 - q) + (1 - \tilde{q})q)](r) \\
&= \tilde{p}p\tilde{q}q + \tilde{p}p(1 - \tilde{q})(1 - q) + (1 - \tilde{p})(1 - p)\tilde{q}q \\
&\quad +(1 - \tilde{p})(1 - p)(1 - \tilde{q})(1 - q) - \tilde{p}(1 - p)\tilde{q}q \\
&\quad -\tilde{p}(1 - p)(1 - \tilde{q})(1 - q) - (1 - \tilde{p})p\tilde{q}q \\
&\quad -(1 - \tilde{p})p(1 - \tilde{q})(1 - q) - \tilde{p}p\tilde{q}(1 - q) + \tilde{p}p(1 - \tilde{q})q \\
&\quad -(1 - \tilde{p})(1 - p)\tilde{q}(1 - q) - (1 - \tilde{p})(1 - p)(1 - \tilde{q})q \\
&\quad \tilde{p}(1 - p)\tilde{q}(1 - q)r + \tilde{p}(1 - p)(1 - \tilde{q})qr \\
&\quad +(1 - \tilde{p})p\tilde{q}(1 - q)r + (1 - \tilde{p})p(1 - \tilde{q})qr \\
&= \tilde{p}\tilde{q}(pq - (1 - p)q - p(1 - q) + (1 - p)(1 - q)r) \\
&\quad \tilde{p}(1 - \tilde{q})(p(1 - q) - (1 - p)(1 - q) - pq + (1 - p)qr) \\
&\quad (1 - \tilde{p})\tilde{q}((1 - p)q - pq - (1 - p)(1 - q) + p(1 - q)r) \\
&\quad (1 - \tilde{p})(1 - \tilde{q})((1 - p)(1 - q) - p(1 - q) - (1 - p)q + pqr) \\
&= \tilde{p}\tilde{q}(\pi_{\boldsymbol{p}^*}^{\boldsymbol{q}^*} + \rho_{\boldsymbol{p}^*}^{\boldsymbol{q}^*}) - \tilde{p}(1 - \tilde{q})(\pi_{\boldsymbol{q}^{@}}^{\boldsymbol{p}^*} + \rho_{\boldsymbol{q}^{@}}^{\boldsymbol{p}^*}) \\
&\quad -(1 - \tilde{p})\tilde{q}(\pi_{\boldsymbol{p}^{@}}^{\boldsymbol{q}^*} + \rho_{\boldsymbol{p}^{@}}^{\boldsymbol{q}^*}) + (1 - \tilde{p})(1 - \tilde{q})(\pi_{\boldsymbol{p}^{@}}^{\boldsymbol{q}^{@}} + \rho_{\boldsymbol{p}^{@}}^{\boldsymbol{q}^{@}}) \\
&= \tilde{p}\tilde{q}v - \tilde{p}(1 - \tilde{q})v - (1 - \tilde{p})\tilde{q}v + (1 - \tilde{p})(1 - \tilde{q})rv \\
&= v(\tilde{p}\tilde{q} - \tilde{p}(1 - \tilde{q}) - (1 - \tilde{p})\tilde{q} + (1 - \tilde{p})(1 - \tilde{q})r) \\
&= v(pq - p(1 - q) - (1 - p)q + (1 - p)(1 - q)r) \\
&= v^2
\end{aligned}$$

where in the third step we used the identity

$$1-(\tilde{p}p+(1-\tilde{p})(1-p)) = 1-\tilde{p}p-1+\tilde{p}+p-\tilde{p}p = \tilde{p}(1-p)+(1-\tilde{p})p$$

the substitutions used in the sixth step come from Appendix B and can be verified with some additional algebra and in the ninth step we used the identities

$$\tilde{p} = p \text{ and } \tilde{q} = q$$