A game of competition over risk

Abstract

This article investigates models of competition where actors choose their level of risk and are rewarded for taking on more risk than the others. These models are highly relevant to real-life situations of competition over prices, such as between firms, banks, and insurance companies seeking to acquire more customers.

Starting from a simple normal form game with two players on a continuous action space, we prove the existence and uniqueness of a Nash equilibrium with an analytical solution. This serves as a foundation to incorporate more realistic features such as multiple players, market frictions and correlations between firm risks. We verify experimentally that the Nash equilibrium is a correlated equilibrium. Finally, we show that variants of regret matching offer performant algorithms to approximate the Nash equilibrium in our game. Our code is open-source [1] and can be applied to any game in normal form.

Our contribution is also methodological and promotes the use of algorithms in economics. Most models are overly simplified so that they admit an analytical solution. However, general algorithmic solvers could handle much more complex models, not having closed form solutions or taking data as input. While the literature has been focused on finite games, we aim to provide the reader with such algorithms in the context of games with a continuous action set. Overall, our approach provides a rigorous framework for modeling and analyzing strategic interactions in continuous action games, with wide-ranging applications in economics, finance, and policy-making.

1. Introduction

Competition over risk is a ubiquitous feature of many real-life situations, ranging from financial markets to environmental policy. In these situations, actors face a trade-off between the potential rewards of taking on more risk

[1] available at [redacted]

and the possibility of negative outcomes, such as bankruptcy or environmental disaster. Understanding the behavior of actors in these situations and predicting outcomes is crucial for policymakers, regulators, and investors.

Game theory provides a powerful framework for analyzing these situations, as it models the strategic interactions between actors and their incentives to take certain actions. In particular, we focus on normal form games, which represent situations where each player chooses a strategy and receives a payoff based on the joint actions of all players. In this article, we study continuous models of competition, where actors can choose the level of risk and are rewarded for taking more risk.

A key concept in game theory is that of a Nash equilibrium, which represents a stable point of the game where no player has an incentive to deviate from their chosen strategy. In a normal form game, a Nash equilibrium is a set of strategies where each player's strategy is a best response to the other players' strategies. Nash equilibria are a fundamental concept in game theory and have been widely used in economics, political science, and other fields to model strategic behavior.

In this article, we start by examining a simple normal form game with two players and prove the existence and uniqueness of a Nash equilibrium with an analytical solution. This simplistic game serves as a backbone that we can improve on by incorporating more realistic features, such as market frictions and correlations between firm risks. We then study several methods that can be used to approximate Nash equilibria in continuous action games and compare their running time and solution quality.

Finally, we apply our findings to the setting of competition between financial actors providing saving products to the public, where risk-taking plays a crucial role in the decision-making process. By studying these models of competition over risk, we hope to provide insights into the behavior of actors in these situations and develop tools for predicting outcomes and improving decision-making.

2. A simple model for competition over risk

2.1. Description

In this section, we introduce a simple model of competition over risk that serves as a backbone for our study. We consider a situation where two actors, denoted as Player 1 and Player 2, engage in competition by taking actions that make them more attractive to customers but also increase their risk of failure. For example, firms may choose to lower their prices to attract more customers but in doing so, they increase the likelihood of not being able to repay their loans. Similarly, insurance companies may lower their premiums to attract more customers but this comes at the cost of a higher risk of failure. Banks may increase their deposit rates to attract more customers but this also increases their vulnerability to liquidity crises.

In our model, each player directly sets their failure probability, denoted as r_p. While this assumption may not be realistic in practice, we note that in many situations, firms use models that map real-world actions, such as setting prices or premiums, to failure probabilities. This mapping is often a monotonous function that can be inverted to yield real-world actions from failure probabilities, making our model practical. Based on the failure probabilities set by each player, the players can randomly "lose" the game. In our simple model, this translates into being applied a penalty, denoted as P. We assume $P > 0$. We assume that the failure events are independent, meaning that each player draws a uniform random variable f_p from the interval $[0, 1]$ and fails if $f_p < r_p$. We will later introduce correlations between the variables f_p to model real-life situations where correlations may be positive or negative.

After the failure events are determined, the players that did not fail compare their risk levels, and the player that played the highest risk level is rewarded with a payoff, denoted as R. We assume that $R > 0$. In the case of ties between risk levels, we consider several ways of resolving them, such as none of the players receiving the reward, the reward being shared equally between them, or the reward being randomly given to one of them.

We note that, as shown later, the optimal strategies in our model are modeled by real distributions, which means that the probability of ties is zero. However, when we use discrete action sets to compute approximate Nash equilibria, the action sets can overlap, and we implement the first two variations of resolving ties (the last two are equivalent in expectation). This simple model serves as a foundation for our study, and we will extend it by introducing correlations between the players' failure probabilities and market frictions in subsequent sections. For two players, assuming $r_1 > r_2$, the outcome matrix will be:

	$f_1 \geq r_1$	$f_1 < r_1$
$f_2 \geq r_2$	$R, 0$	$-P, R$
$f_2 < r_2$	$R, -P$	$-P, -P$

Each cell contains the rewards to each player. For example, in the upper left cell, no failure happens. Since we assumed $r_1 > r_2$, player 1 gets R and player 2 gets 0.

2.2. Equivalence to a normal-form game

We can represent the game described above in the framework of extensive-form games by modeling the drawing of the random variables f_p using Chance nodes. Since the outcomes are subject to randomness, it is natural to assume that the actors operate under the expected utility hypothesis, which implies that they possess a von Neumann–Morgenstern utility function. Consequently, we can define a normal-form game with payoffs equal to the expected payoffs of the corresponding extensive-form game. By doing so, we can leverage the theory of normal-form games and apply various solution concepts, such as Nash equilibria, to analyze the competition between the actors.

Proposition 2.1. *The expected utilities u_p are computed as follows in the 2-player game:*

$$u_2(r_1, r_2) = u_1(r_2, r_1) \text{ (symmetry)}$$
$$u_1(r_1, r_2) = r_2(1 - r_1)R - r_1 P + [r_1 > r_2](1 - r_1)(1 - r_2)R$$

where $[\cdot]$ is the Iverson bracket.

Proof. Player 1 can fail with probability r_1, in which case they lose P. If Player 2 loses and Player 1 does not, which happens with probability $r_2(1 - r_1)$, Player 1 wins R. Finally, if none of the players fails, when $r_1 > r_2$, Player 1 can win R. □

It is possible to encompass the shared payoff in case of ties by defining the Iverson bracket to be $\frac{1}{2}$ when $r_1 = r_2$. Figure 1 shows what the reward function of Player 1 looks like when Player 2 adopts the fixed strategy $r_2 = 0.2$.

The discontinuity of our game is similar to two games: the War of Attrition game from Smith (1974) and the visibility game from Lotker et al. (2008). In the War of Attrition game, each player independently chooses a time to quit the game. The player who stays in the game for the longest time wins a prize. However, both players incur a cost that increases over time while they are still in the game. In the visibility game, the payoff of

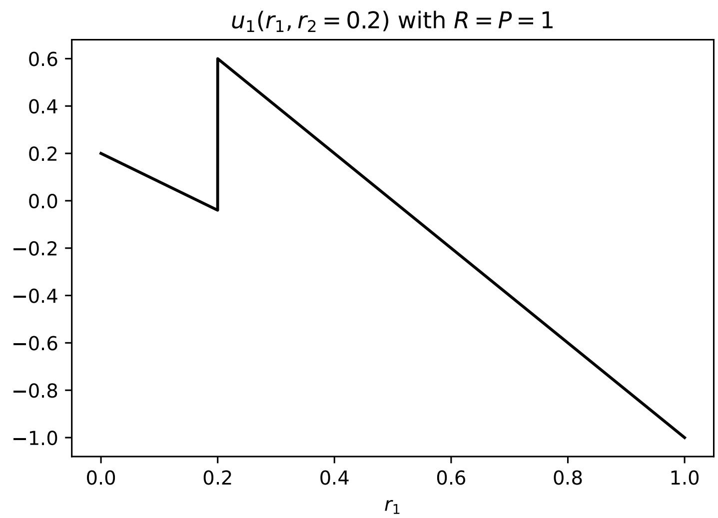


Figure 1: Reward function

each player is the difference with the next player, or 1 for the player that plays the largest move. A major difference between our game and those two games is that we model the probability of failure. This means, for example, that the player taking less risk can still win the reward if the first player fails.

However, the structure of our problem and the analytical solution of the Nash equilibrium are similar to Lotker et al. (2008). We name our game the Competition over Risk game and will write it CoR in the rest of the article.

2.3. Nash equilibrium

A Nash equilibrium is a set of strategies, one for each player, such that no player can improve their payoff by unilaterally changing their strategy, given the strategies of the other players. In other words, each player's strategy is the best response to the strategies chosen by the other players. Nash equilibria are important because they provide a way to predict the outcome of a game if each player acts rationally and selfishly. They can also help explain why certain outcomes occur in real-world situations.

In our model of competition over risk, finding Nash equilibria can help us understand how firms, banks, and insurance companies behave when they compete over prices and take different levels of risk. By analyzing the Nash equilibria of our model, we can predict how different players will act and what the resulting outcomes will be. Moreover, we can compare the efficiency of different equilibria and use them as a benchmark to evaluate the performance of different strategies.

As in the game of Lotker et al. (2008), we can prove that there is no pure Nash equilibrium, that is, a deterministic optimal strategy.

Theorem 2.2. *The CoR game does not admit any pure Nash equilibrium.*

Proof. Suppose the existence of an equilibrium s_1, s_2. Suppose that $s_1 > s_2$. Then Player 1 can improve their payoff by playing $s_1 - \varepsilon$ since they still get the reward and take less risk. By symmetry, this implies that $s_1 = s_2$. If $s_1 < 1$, then player 1 can improve their situation by playing $s_1 + \varepsilon$ since they get R (or $\frac{R}{2}$ if the reward is shared). If $s_1 = 1$ then the payoff is $-P < 0$ with probability 1 so it is better to play 0 which gives payoff 0 with probability 1. □

Definition 2.1. A strategy s (a couple of strategies) is Pareto optimal if there is no other strategy s' such that $\forall p, u_p(s) \leq u_p(s')$ and $\exists p, u_p(s) < u_p(s')$. It is ε-Pareto optimal if there is no strategy s' such that $\forall p, u_p(s) \leq u_p(s')$ and $\exists p, u_p(s) + \varepsilon < u_p(s')$.

Remark. *If the reward is shared in case of tie, the pure strategy* $(0, 0)$ *gives reward* $\frac{R}{2}$ *to each player. This strategy is Pareto-optimal.*

Theorem 2.3. *For every* ε*, there is a* ε*-Pareto optimal strategy that gives* $\frac{R-\varepsilon}{2}$ *to each player.*

Proof. Let us consider the joint mixed strategy where each player plays uniformly at random in the interval $[0, 2\varepsilon]$. The payoff is

$$\begin{aligned}
\mathbb{E}[u_1] &= \mathbb{E}\left[r_2(1-r_1)R - r_1 P + [r_1 > r_2](1-r_1)(1-r_2)R\right] \\
&= \varepsilon(1-\varepsilon)R - \varepsilon P + \frac{(1-\varepsilon)^2}{2} R \\
&\to_{\varepsilon \to 0} \frac{R}{2}
\end{aligned}$$

so by taking ε small enough we can get as close to $\frac{R}{2}$ as we want.

If each player gets payoff $\frac{R-\varepsilon}{2}$ then no player can get ε without degrading the other's performance else the total payoff would be more than R. □

However, the ε-Pareto strategy is highly concentrated around 0, incentivizing players to deviate and increase their chances of winning R without taking on additional risk. Thus, this strategy fails to form a Nash equilibrium.

Fortunately, the CoR game possesses a unique Nash equilibrium, a powerful property that showcases the strength of our approach. Moreover, this equilibrium is symmetric.

For finite games, Nash himself proved the existence of mixed Nash equilibria (Nash, 1950), while Glicksberg's theorem extended this result to continuous reward functions (Glicksberg, 1952). Dasgupta and Maskin (Dasgupta and Maskin, 1986) established conditions under which discontinuous games can possess Nash equilibria and symmetric games can admit symmetric equilibria.

The uniqueness of the Nash equilibrium is a highly desirable property, with most models using concave reward functions to ensure it. Therefore, it is noteworthy that the CoR game exhibits a unique Nash equilibrium.

We recall Theorem 2.1 from Lotker et al. (2008):

Theorem 2.4. *Let $(f_1, \ldots, f_n)$ be a Nash equilibrium point, with expected payoff u_i^* to Player i at the equilibrium point. Let $u_i(x)$ (as an abuse of notation) denote the expected payoff for Player i when he plays the pure strategy x and all other players play their equilibrium mixed strategy. Then $u_i(x) \leq u_i^*$ for all $x \in [0, 1]$, and furthermore, there exists a set $\mathcal{Z}$ of measure 0 such that $u_i(x) = u_i^*$ for all $x \in \mathit{support}(f_i) \setminus \mathcal{Z}$.*

This theorem means that at the Nash equilibrium, almost any move that is in the support of a player's strategy should give them the same (maximal) payoff. This theorem is crucial to find the equilibrium in the CoR game.

Theorem 2.5. *Up to a set of measure zero, the CoR game admits a unique Nash equilibrium. This equilibrium is symmetric and its distribution is $f(x) = \left[x < 1 - \sqrt{\frac{k-1}{k+1}}\right] \frac{k-1}{(1-x)^3}$ with $C := \frac{P}{R}$ and $k := \sqrt{(C+1)^2 + 1}$. The the average move is $\bar{r} = k - (C+1)$ and the utility of each player is $u^* = R\bar{r}$.*

Proof. See Appendix A for a full proof. For a less rigorous treatment, refer to the proof of the more general Theorem 3.1. $\square$

At $C = 1$, the cutoff value is $1 - \sqrt{\frac{\sqrt{5}-1}{\sqrt{5}+1}} = 2 - \phi \approx 0.382$ with ϕ the Golden ratio. We plot the distribution in Figure 2.

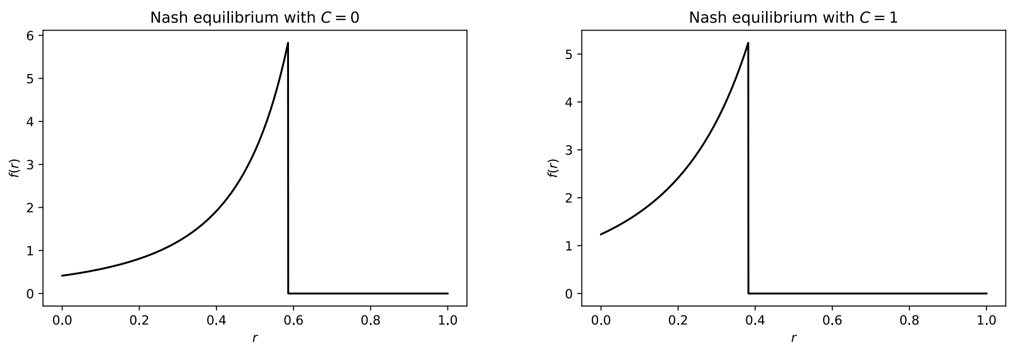


Figure 2: Nash equilibrium

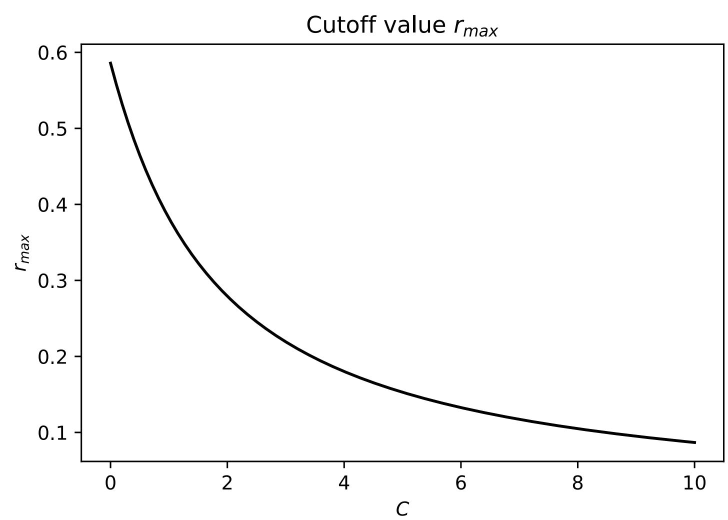


Figure 3: The cutoff goes to zero when $C \to \infty$.

The behavior of the cutoff r_{max} is displayed in Figure 3. Unsurprisingly, when $C \to \infty$, the penalty P becomes much larger than the reward R and the players play closer to 0.

The case when $C \to 0$ is more surprising: the maximal cutoff value at $C = 0$ is $h = 1 - \sqrt{\frac{\sqrt{2}}{\sqrt{2}+2}} \approx 0.356$. This is because even if the penalty is 0, the players cannot get the reward if they "lose", which prevents them from taking too much risk. We plot the distribution in Figure 2. See Appendix Appendix B for more figures.

3. Generalization to multiple players

Quite naturally, we wonder what the Nash equilibrium looks like for multiple players. The visibility game of Lotker et al. (2008) probably does not admit an analytical solution and they instead give an algorithm to produce approximate solutions. We show that the CoR game for multiple players admits a unique symmetric equilibrium and present a new numerical algorithm to compute it. We finally study the asymptotic behavior of the equilibrium.

3.1. Nash equilibrium

Interestingly, our Correlation over Risk game admits an analytical solution even for multiple players. More precisely:

Theorem 3.1. *There is a unique symmetric Nash equilibrium for in the CoR game with n players defined by*

$$f(x) = \frac{C + w}{(n-1)(1-x)^{2+\frac{1}{n-1}}(Cx + w)^{1-\frac{1}{n-1}}}$$

for some constants r_{max} and $w := \bar{r}^{n-1}$ (the probability of winning when taking no risk) such that

$$\int_0^{r_{max}} f(x)dx = 1$$
$$\int_0^{r_{max}} xf(x)dx = \bar{r}$$

Proof. We adapt the proof of Theorem 2.5 and start by assuming the existence of a symmetric mixed equilibrium defined by the probability density f.

First we derive a nice expression for $u(x)$, defined as the utility of one player choosing move x while the others play according to f. For all $x \in support(f)$:

$$u(x) = -xP + (1-x)\left(\int_0^x f(y)dy + \int_x^1 yf(y)dy\right)^{n-1} R$$

This equation is quite natural: the player loses P with probability x. If they survive, with probability $1-x$, they need the $n-1$ other players to either play a lower value or play a higher value and fail. We can suppose as previously that 0 is in the support to subtract $u(0)$. We write $\bar{r}$ for the expectation of the action r under f.

$$\left(\frac{\bar{r}^{n-1} + xC}{1-x}\right)^{\frac{1}{n-1}} = \int_0^x f(y)dy + \int_x^1 yf(y)dy$$

We define $w := \bar{r}^{n-1}$ to be the probability of winning when taking no risk, we derivate and divide by $1-x$ to obtain:

$$f(x) = \frac{C+w}{(n-1)(1-x)^{2+\frac{1}{n-1}}(Cx+w)^{1-\frac{1}{n-1}}}$$

Finally we can solve $\int_0^{r_{max}} f(x)dx = 1$ and $\int_0^{r_{max}} xf(x)dx = \bar{r}$. We relegate the description of the numerical estimation of r_{max} and w to Appendix C. $\square$

We display the behavior of the solution for multiple players in Figure 4.

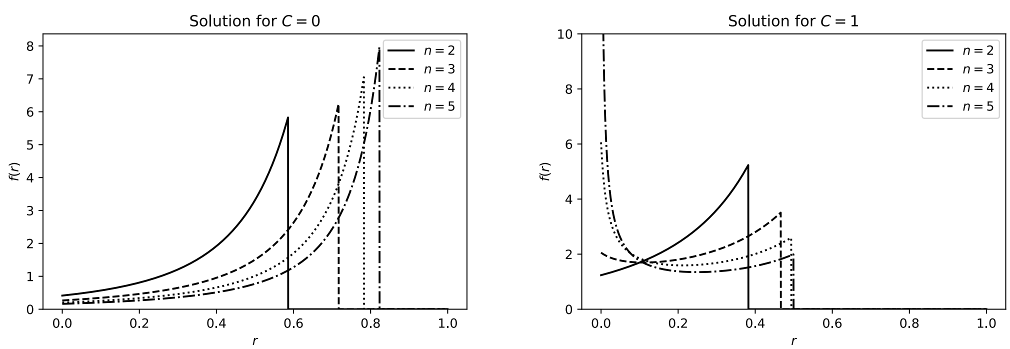


Figure 4: We observe a clear difference between the cases $C = 0$ (no penalty) and $C = 1$ (presence of a penalty). In both cases, the cutoff increases. However, the average risk seems to decrease sharply when there is a nonzero penalty with a mode at $r = 0$.

3.2. Asymptotic behavior

We are interested in studying the equilibrium when the number of players goes to infinity. For fixed C, we have the following:

Proposition 3.2. *When $n \to \infty$, $\lim r_{max} = \frac{1}{1+C}$ and $\bar{r} \sim \frac{1}{nC}$*

Proof. We verify experimentally that r_{max} is never close to 0 or 1 and that $\bar{r} \to 0$. Equation C.1 gives

$$\frac{w + nC(1 - r_{max}) + Cr_{max}}{n(1 - r_{max})(C + w)} \sqrt[n-1]{\frac{Cr_{max} + w}{1 - r_{max}}} = 1 + \frac{w + nC}{n(C + w)}\bar{r}$$

$$\sqrt[n-1]{\frac{Cr_{max} + w}{1 - r_{max}}} \to 1$$

Equation C.2 gives

$$\frac{w - nw(1 - r_{max}) + Cr_{max}}{n(1 - r_{max})(C + w)} \sqrt[n-1]{\frac{Cr_{max} + w}{1 - r_{max}}} = \bar{r}\frac{w + nC}{n(C + w)}$$

$$\frac{w}{C} + \frac{r_{max}}{n(1 - r_{max})} \sim \frac{r_{max}}{n(1 - r_{max})} \sim \bar{r}$$

using $w = \bar{r}^{n-1} = o(\bar{r})$.

$\sqrt[n-1]{\frac{Cr_{max}+w}{1-r_{max}}} \to 1$ implies $\frac{r_{max}}{1-r_{max}} \to \frac{1}{C}$ and $r_{max} \to \frac{1}{1+C}$.

Finally, $\bar{r} \sim \frac{1}{nC}$. □

We illustrate this behavior in Figure 5.

A common concept in game theory is the price of anarchy PoA (Koutsoupias and Papadimitriou, 1999). The price of anarchy is the ratio between the Pareto optimum and the Nash equilibrium. It is easy to generalize Theorem 2.3 for multiple players and show that the reward can be split almost perfectly to obtain a Pareto optimal utility $\frac{R}{n}$. The utility of our symmetric equilibrium is $R\bar{r}^{n-1} = R\bar{r}^{n-1} = Rw$. Hence, $PoA = 1/nw$. We will instead compute the efficiency $E = \frac{1}{PoA} = nw \in [0, 1]$.

Experimental result 3.3. When $C = \frac{1}{n^e}$ with $e \geq 0$, the efficiency $E = nw$ of the Nash equilibrium goes to 0 if $e \leq 1$ and it goes to 1 if $e > 1$. We plot the behavior of E in Figure 6.

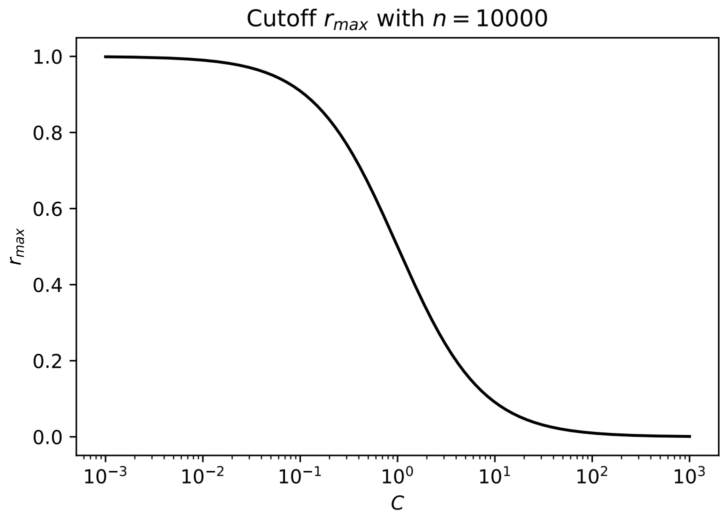


Figure 5: The plot of r_{max} to $\log(C)$ is very similar to the function $\frac{1}{1+\exp(\cdot)}$. This is because $r_{max} \sim \frac{1}{1+C}$.

We interpret Result 3.3 as an indication that resources, here modeled by the ratio $\frac{1}{C} = \frac{R}{P}$ of rewards to penalties, need to scale faster than the number of players for them to adopt an efficient behavior. Scarcity of resources creates an inefficient Nash equilibrium.

4. Extensions of the Competition over Risk Game

4.1. Market frictions

One limitation of our model is the assumption that the utility functions are discontinuous at a certain threshold. While this is appropriate for certain scenarios such as call for bids, it may not hold in other real-life situations that involve noisy evaluations or aggregate many individual choices. To address this limitation, we propose replacing the threshold $[r_1 > r_2]$ with a smooth choice model using the logistic function, $\sigma_\tau(r_1 - r_2)$, where σ_τ is the scaled sigmoid:

$$\sigma_\tau(x) := \frac{1}{1 + \exp\left(-\frac{x}{\tau}\right)}$$

Recall that the failure events are $f_p < r_p$. For a game between two players, the outcome matrix can be represented as follows:

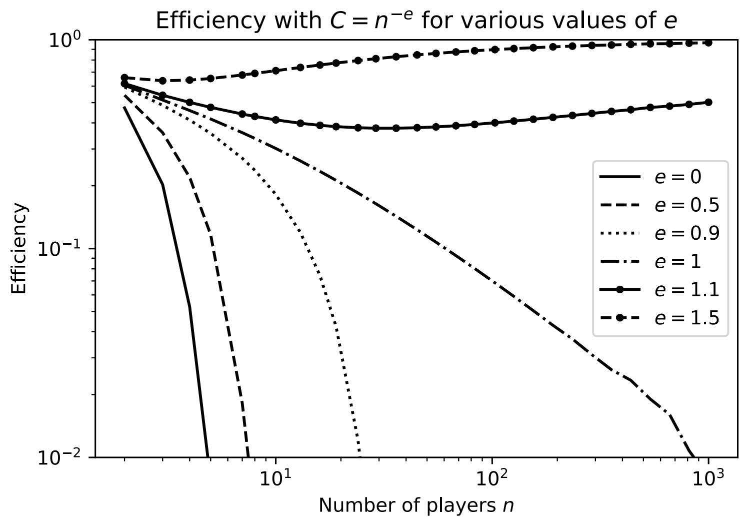


Figure 6: The efficiency clearly goes to 0 even when $e = 1$. When $e > 1$, it seems that $E \to 1$. Values for greater values of n suffer of numerical precision issues as $\log w \to 0$.

	$f_1 \geq r_1$	$f_1 < r_1$
$f_2 \geq r_2$	$R\ \sigma_\tau(r_1 - r_2), R\ \sigma_\tau(r_2 - r_1)$	$-P, R$
$f_2 < r_2$	$R, -P$	$-P, -P$

Proposition 4.1. *The expected utilities u_p for the 2-player game with frictions are computed as follows:*

$$u_2(r_1, r_2) = u_1(r_2, r_1) \text{ (symmetry)}$$
$$u_1(r_1, r_2) = r_2(1 - r_1)R - r_1 P + (1 - r_1)(1 - r_2)\sigma_\tau(r_1 - r_2)R$$

As $\tau \to 0$, σ_τ approaches the Heaviside step function and market frictions disappear.

4.2. Correlation between risks

In the real world, risks are often correlated, which is not accounted for in our current model. To incorporate correlation between risks, we can introduce joint distributions for the failure events f_p, which occur according to latent variables.

In our model, we assume that f_p follows a uniform distribution. To introduce correlation between f_1 and f_2, we use the well-known NORTA (NORmal

To Anything) method (Cario and Nelson, 1997). This method allows us to create a joint distribution (f_1, f_2) such that the marginals are uniform distributions and the Pearson correlation between f_1 and f_2 can be set to any arbitrary value.

Following NORTA, we define $f_p = \Phi(z_p)$, where Φ is the cumulative distribution function of the Normal distribution, and

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim \mathcal{N}(\mu, \Sigma)$$

with $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\Sigma = \begin{pmatrix} 1 & \rho(z_1, z_2) \\ \rho(z_1, z_2) & 1 \end{pmatrix}$. Here, ρ is the Pearson correlation coefficient between z_1 and z_2, which determines the correlation between f_1 and f_2. As shown in Cario and Nelson (1997), specifying a correlation between z_p or f_p is equivalent to specifying $\rho(f_1, f_2)$. Specifically, we have

$$\rho(f_1, f_2) = \frac{6}{\pi} \sin^{-1}\left(\frac{\rho(z_1, z_2)}{2}\right)$$

Hence, we use ρ to denote $\rho(z_1, z_2)$ throughout the rest of the document. This model is well-suited to real-world scenarios, such as financial portfolios, where z_p can represent the returns on investments. In such cases, joint distributions of portfolios are typically modeled as multivariate normal distributions, and r_p corresponds to the Value at Risk $v_p = \Phi^{-1}(r_p)$ through the bijective function Φ, such that the failure event $z_p < v_p$ is equivalent to $f_p < r_p$.

Proposition 4.2. *The expected utilities u_p for the 2-player game with frictions and correlated risks are computed as follows:*

$$u_2(r_1, r_2) = u_1(r_2, r_1) \text{ (symmetry)} \tag{1}$$
$$u_1(r_1, r_2) = (r_2 - \tilde{r})R - r_1 P + (1 - r_1 - r_2 + \tilde{r})\sigma_\tau(r_1 - r_2)R \tag{2}$$

where $\tilde{r} := \Phi_\rho(\Phi^{-1}(r_1), \Phi^{-1}(r_2))$ is the probability of joint failure, with

$$\Phi_\rho(v_1, v_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{v_1} \int_{-\infty}^{v_2} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho)^2}\right) dy\, dx$$

the cumulative distribution of the bivariate normal distribution with correlation ρ.

In the absence of noise, when $\rho = \pm 1$, it is also possible to calculate the Nash equilibrium analytically:

Theorem 4.3. *For $\rho = 1$, the equilibrium is given by:*

$$p(x) = \frac{1+C}{1-x}\left[x < 1 - \exp\left(-\frac{1}{C+1}\right)\right]$$

We have

$$\bar{r} = 1 - (C+1)\left(1 - \exp\left(-\frac{1}{C+1}\right)\right)$$

For $\rho = -1$, the equilibrium is given by:

$$p(x) = \frac{C}{(1-2x)^{3/2}}\left[x < \frac{1}{2} - \frac{C^2}{2(C+1)^2}\right]$$

We have

$$\bar{r} = \frac{1}{2C+2}$$

Proof. See Appendix D. □

5. Computing approximate Nash equilibrium

5.1. Approximations to games and equilibria

In this section, we define some key concepts and metrics related to games and equilibria.

For a given game with n players, we use $u_i(\sigma)$ to denote the reward of player i when all players follow the strategy $\sigma = (\sigma_1, \ldots, \sigma_n)$[2]. A strategy σ is said to be a Nash equilibrium if it satisfies the following condition for all players i and all alternative strategies $\sigma'_i \in \Sigma_i$:

$$u_i(\sigma) \geq u_i(\sigma'_i, \sigma_{-i})$$

where σ_{-i} is the strategy of all players but i, and Σ_i is the set of actions available to player i.

A game is said to be continuous if the action space Σ_i is compact and u_i is continuous. In such games, it is possible to approximate the Nash equilibria using a sequence of games over a reduced finite support, which leads

[2]This σ is not the same as the scaled sigmoid σ_τ defined earlier.

to Glicksberg's theorem, without relying on Kakutani's theorem (Myerson, 1997).

In our CoR game, which has a few points of discontinuity, it is also possible to approximate the Nash equilibria using a similar method. However, we do not provide a proof of this here, as the introduction of frictions makes our game continuous anyway.

To measure the closeness of a strategy σ to a Nash equilibrium, we use the NashConv metric (Lanctot et al., 2017):

$$\textsc{NashConv}(\sigma) = \sum_{i=1}^{n} \max_{s_i \in \Sigma_i} u_i(s_i, \sigma_{-i}) - u_i(\sigma)$$

Note that this metric only considers pure strategies $s_i \in \Sigma_i$, due to the linearity of the payoff function for mixed strategies.

The $\textsc{NashConv}$ metric satisfies $\textsc{NashConv}(\sigma) \geq 0$, with equality holding only for a Nash equilibrium. This implies that $\textsc{NashConv}(\sigma)$ corresponds to the notion of ε-Nash equilibrium, where a ε-Nash equilibrium σ has $\textsc{NashConv}(\sigma) = n\varepsilon$.

For a finite action space, $\textsc{NashConv}$ is easy to compute since Σ_i is finite. However, for a continuous action space, no such metric is known. Nonetheless, we can approximate $\textsc{NashConv}$ by taking the maximum over a finite sample of points from Σ_i. This sample can be chosen randomly, or if Σ_i is an interval or a product of intervals of $\mathbb{R}$, we can use a grid.

In our CoR game, the action space is $[0, 1]$. Here, we use quasi-random numbers to measure the closeness to a Nash equilibrium, inspired by the literature on hyperparameter sampling (Bousquet et al., 2017) and the efficiency of quasi-Monte-Carlo methods (Sobol', 1990). Specifically, we define the QuasiNashConv metric as:

$$\textsc{QuasiNashConv}(\sigma, m) = \sum_{i=1}^{n} \max_{s_i \in \textsc{Sobol}(m)} u_i(s_i, \sigma_{-i}) - u_i(\sigma)$$

where $\textsc{Sobol}(m)$ is a set of m quasi random numbers drawn using Sobol's method (Sobol', 1967).

5.2. Correlated Equilibria

A Nash equilibrium is a set of strategies where no player can improve their payoff by unilaterally changing their strategy, assuming that all other players' strategies remain unchanged. However, in some games, players may benefit from coordinating their actions in ways not captured by traditional Nash equilibrium. This is where the concept of correlated equilibrium comes in.

A correlated Nash equilibrium is a set of correlated strategies where no player can improve their expected payoff by unilaterally changing their strategy, given that they observe the correlation signal. This correlation signal is not necessarily a message or communication between the players, but rather a shared random variable that affects each player's strategy consistently.

Definition 5.1. A correlated Nash equilibrium is a joint distribution σ over all moves $\Sigma_1 \times \Sigma_2 \times \ldots \times \Sigma_n$ such that for any player i and any strategy modification $\phi : \Sigma_i \to \Sigma_i$,

$$u_i(\sigma_i, \sigma_{-i}) \geq u_i(\phi(\sigma_i), \sigma_{-i})$$

Thus, a Nash equilibrium can be viewed as a correlated Nash equilibrium that can be decomposed into independent strategies for each player. It is evident that any Nash equilibrium is a correlated Nash equilibrium.

Correlated equilibria are more suitable for the real world because they allow for a broader range of possible outcomes that can arise through coordination among the players, without necessarily requiring communication or binding agreements between them.

In many real-world scenarios, it is challenging or impossible for players to communicate and make binding agreements, or they may not have complete information about the strategies of the other players. Correlated equilibria provide a way for players to achieve coordination and cooperation without requiring such communication or information, by relying on shared random variables that affect each player's strategies consistently.

Finally, correlated equilibria can also capture situations where players have some degree of trust or social norms that encourage them to coordinate their actions in a specific way. For instance, in a repeated game where players interact with each other over a long period, they may develop a sense of reciprocity or reputation that encourages them to follow a certain coordinated strategy.

5.3. Finding Correlated Equilibria with Linear Solvers

Correlated equilibria are of interest because they can be computed more easily for a finite action set.

A joint strategy can be represented by a mapping of probabilities:

$$\mathrm{Pr}_\sigma(s_1, s_2, \ldots, s_n) := \mathrm{Pr}[\sigma = (s_1, s_2, \ldots, s_n)]$$

for all joint actions $(s_1, s_2, \ldots, s_n)$. Therefore, the equation from Definition 5.1 is linear in these probabilities. An additional equation is that probabilities must sum to 1, and all probabilities are constrained to be positive. For two players, the equations are:

$$\forall (s_1, s_1'), \sum_{s_2} \mathrm{Pr}_\sigma(s_1, s_2) u_1(s_1, s_2) \geq \sum_{s_2} \mathrm{Pr}_\sigma(s_1, s_2) u_1(s_1', s_2)$$

$$\forall (s_2, s_2'), \sum_{s_1} \mathrm{Pr}_\sigma(s_1, s_2) u_1(s_1, s_2) \geq \sum_{s_1} \mathrm{Pr}_\sigma(s_1, s_2) u_1(s_1, s_2')$$

$$\forall (s_1, s_2), \mathrm{Pr}_\sigma(s_1, s_2) \geq 0$$

$$\sum_{s_1, s_2} \mathrm{Pr}_\sigma(s_1, s_2) = 1$$

The set of correlated equilibria is thus a convex polytope P. It is possible to find the boundary in any direction using a linear programming solver.

It is also possible to check that the correlated equilibrium is unique and is a Nash equilibrium by trying to maximize and minimize each variable over the polytope. If the maximum and minimum are equal for each variable, then the polytope only contains one point. Another method described in Appa (2002) checks the uniqueness of a solution to a linear program by solving a new linear program. However, that method requires a reformulation of the linear program as $\max cx$ s.t. $Ax = b, x \geq 0$, which is cumbersome in our case. We propose a simple randomized method (algorithm 5.3) that can produce confidence intervals for any confidence level (or p-value).

Theorem 5.1. *Given a polytope P defined by constraints $c_1, \ldots, c_m$*

$$\mathrm{Pr}\left[\textsc{SumDiamSquared}(K, c_1, \ldots, c_m) < \varepsilon\right] \leq F_{\chi^2}\left(\frac{\varepsilon}{diam(P)}, K\right)$$

with $F_{\chi^2}(\cdot, K)$ the cumulative distribution function of the χ^2 distribution with K degrees of freedom.

Algorithm 1 Confidence interval on $diam(P)$

Input: Iterations K, constraints $c_1, \dots, c_m$ defining a polytope P in $\mathbb{R}^n$

```
function SumDiamSquared(K, c_1, ..., c_m)
    for i ← 1, ..., K do
        Sample v_j ~ N(0,1) for j = 1, ..., n
        a_i ← LinProg(v, c)                          ▷ min_{x∈P} v · x
        b_i ← LinProg(−v, c)                         ▷ max_{x∈P} v · x
        d_i ← b_i − a_i
    end for
    return Σ_i d_i^2
end function
```

Input: p-value p

```
function MaxDiameter(p, K, c_1, ..., c_m)
    ε ← SumDiamSquared(K, c_1, ..., c_m)
    q ← Chi2.ppf(p, K)
    d ← ε/q
    return d
end function
```

We used the HiGHS solver (Huangfu and Hall, 2018) to solve the linear optimization subproblems (calls to LinProg). In numerical experiments, we use $K = 5$, confidence $p = 0.95$, and report

$$d_{max} := \textsc{MaxDiameter}(p, K, c_1, \dots, c_m) = \frac{\textsc{SumDiamSquared}(K, c_1, \dots, c_m)}{Q_{\chi^2}(1-p, K)}$$

where $Q_{\chi^2}(\cdot, K)$ is the quantile function of the χ^2 distribution with K degrees of freedom. When the polytope describe probability distributions, we have the bound $d_{max} \leq 2$.

Finally, we make the following trivial remark:

Proposition 5.2. *A correlated equilibrium σ is a Nash equilibrium iff the matrix $(\Pr_\sigma(i,j))_{i,j}$ has rank 1.*

This gives us another numerical method to check that a correlated equilibrium is a Nash equilibrium: compute the second highest eigenvalue and check that it is 0. In numerical experiments, we define λ_1 and λ_2 as the highest and second highest eigenvalues and report the value

$$\lambda := \frac{\lambda_1}{\lambda_2}$$

6. Results

We approximate the game with finite approximations over a grid. Because of the discontinuity of the utility function, we distinguish 2 settings according to a boolean `shift`. In both settings, the two players are proposed a finite number of actions evenly spaced in $[0, 1]$. If `shift = false`, the two players get the same action set. If `shift = true`, the two players get action sets that don't intersect and the actions in $[0, 1]$ are attributed alternatively to each player.

The computational bottleneck of equation 1 is the bivariate normal probability that we compute using a Numba (Lam et al., 2015) implementation of the numerical method of Genz (2004).

Experimental result 6.1. In finite approximations of the CoR game, the Nash equilibrium can be obtained by computing a correlated Nash equilibrium. Furthermore, either the correlated equilibrium is unique or the correlated equilibrium maximizing the total reward is a Nash equilibrium.

We plot d_{max} and λ computed according to section 5.3 for different values of C, τ, ρ and different numbers of actions in Figure 7. Those results are obtained with `shift = true`. We plot the results with `shift = false` in Appendix F.

This motivates the use of algorithms that find correlated equilibria to study our game. In particular, regret matching is a learning strategy that aims to converge to a correlated equilibrium (Hart and Mas-Colell, 1997). At each step, the algorithm provides a distribution of moves for each player. In contrast, Counterfactual Regret Minimization (CFR) (Neller and Lanctot, 2013) is a deterministic variant that accumulates expected regrets instead of playing the game. Although CFR is typically used in extensive form games, we apply it in our normal form game as a deterministic version of Regret Matching. A variant of regret matching called stochastic fictitious play (Fudenberg and Kreps, 1993) involves the softmax function to compute probability distributions. We also test this variant with CFR.

We implement and evaluate all algorithms to solve a discrete version of our game where the action space is reduced to a grid, with and without a

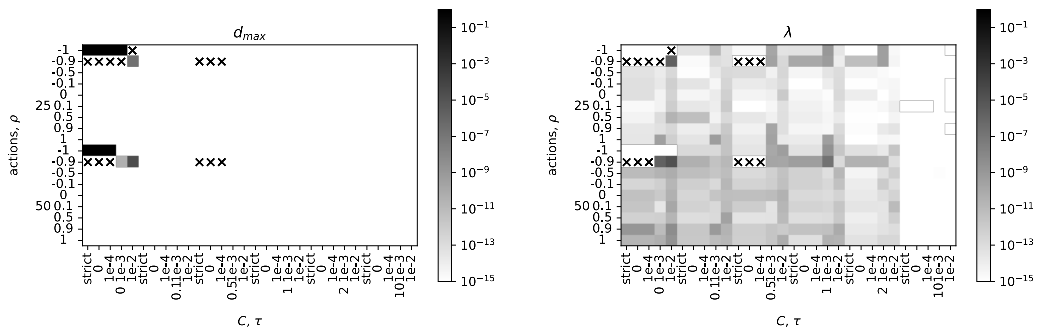


Figure 7: We observe that for almost all values of the parameters, d_{max} is numerically zero. λ is computed for a correlated equilibrium that maximizes the total reward. We cross out cases where the solver failed to find a solution, most probably due to rounding errors which caused the small solution set to disappear. In the few cases where d_{max} was not observed to be zero, λ is clearly zero, which shows that the best correlated equilibrium is a Nash equilibrium. We also interpret the nonzero values of d_{max}

shift. We then report the value of NashConv QuasiNashConv as defined in Section 5.1.

We observe on Figure 8 that vanilla Regret Matching outperforms all methods. In general, Regret Matching performs much better than CFR, even when using less actions, using softmax is detrimental, and shifting the action space does not seem to impact performance much. CFR runs faster as it does not involve any random sampling. Therefore, we use vanilla Regret Matching with shifting in the rest of the experiments for the sake of simplicity. Figure 9 shows that the number of sampled actions is the main factor that drives the quality of solutions.

Experimental result 6.2. Penalties C decrease the average risk taken $\bar{r}$. Market frictions τ decrease the average risk taken and increase the total reward u. Market frictions have more impact on the total reward in high-penalty environments. In particularly inefficient markets with high τ, increasing penalties can counterintuitively improve cooperation and the total reward.

We illustrate that behavior in Figure 10.

Experimental result 6.3. Players take more risks in environments of negative correlation. Negative correlation improves their payoff compared to an absence of correlation.

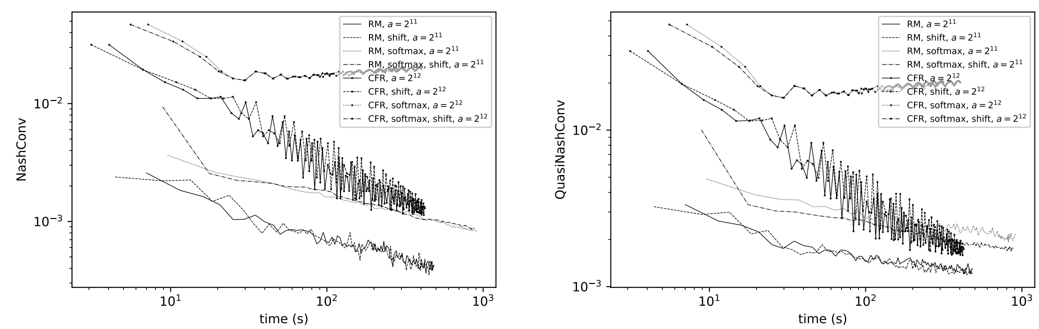


Figure 8: Evaluation of different algorithms to find Nash equilibrium in the standard setting $P = R = 1, \tau = 0, \rho = 0$ without sharing. The game was reduced to a grid of $a = 2^{12}$ actions and QuasiNashConv was evaluated on 2^{15} points. The algorithms ran for 10^4 iterations. For CFR, an iteration is an update to the strategy. For regret matching, we define an iteration as a steps of the sampling, play and update process, so that both CFR and RM do $\mathcal{O}(a^2)$ operations per iteration. We do not include the computation of the utility matrix in the computation time.

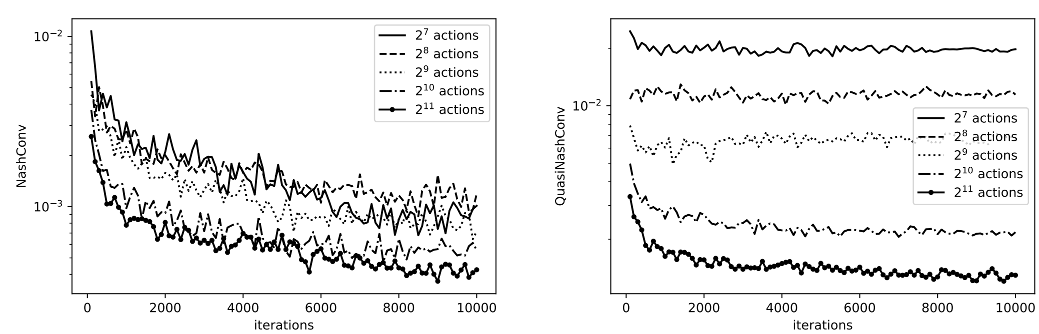


Figure 9: Performance of the vanilla regret matching algorithm to find Nash equilibrium in the standard setting $P = R = 1, \tau = 0, \rho = 0$ without sharing. The quality of the resulting equilibrium depends mainly on the number of actions a. For t iterations, the time complexity of the algorithm is $\mathcal{O}(ta^2)$ and the memory complexity is $\mathcal{O}(a^2)$ as it requires to compute the reward matrix. QuasiNashConv was evaluated on $8a$ points.

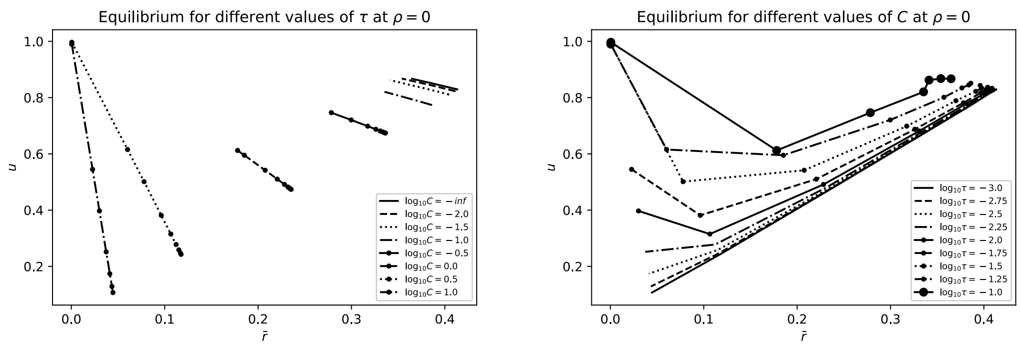


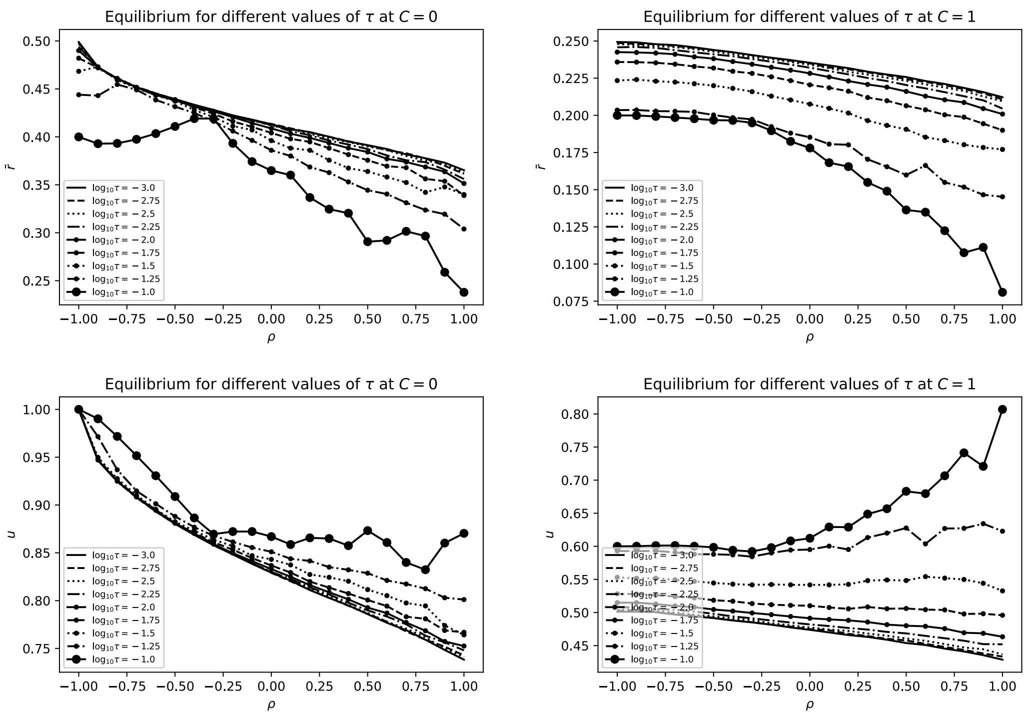
Figure 10: The two figures represent the same points with different level lines: constant C and constant τ. At constant C we observe a linear relationship between u and $\bar{r}$ when changing τ. We also observed that behavior for all values of ρ.

Players take less risks in environments of positive correlation. The impact on performance is negative in efficient markets but can become positive in noisy markets.

We illustrate that behavior in Figure 11.

Acknowledgement

We thank [redacted], [redacted], [redacted] and [redacted] for helpful discussions and useful comments.



Equilibrium for different values of τ at $C = 0$
Equilibrium for different values of τ at $C = 1$
$\bar{r}$
u
ρ
$\log_{10}\tau = -3.0$
$\log_{10}\tau = -2.75$
$\log_{10}\tau = -2.5$
$\log_{10}\tau = -2.25$
$\log_{10}\tau = -2.0$
$\log_{10}\tau = -1.75$
$\log_{10}\tau = -1.5$
$\log_{10}\tau = -1.25$
$\log_{10}\tau = -1.0$

Figure 11:

References

Appa, G., 2002. On the uniqueness of solutions to linear programs. Journal of the Operational Research Society 53, 1127–1132.

Bousquet, O., Gelly, S., Kurach, K., Teytaud, O., Vincent, D., 2017. Critical hyper-parameters: No random, no cry. arXiv preprint arXiv:1706.03200 .

Cario, M.C., Nelson, B.L., 1997. Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix .

Dasgupta, P., Maskin, E., 1986. The existence of equilibrium in discontinuous economic games, i: Theory.

Fudenberg, D., Kreps, D.M., 1993. Learning mixed equilibria. Games and economic behavior 5, 320–367.

Genz, A., 2004. Numerical computation of rectangular bivariate and trivariate normal and t probabilities. Statistics and Computing 14, 251–260.

Glicksberg, I.L., 1952. A further generalization of the kakutani fixed theorem, with application to nash equilibrium points. Proceedings of the American Mathematical Society 3, 170–174.

Hart, S., Mas-Colell, A., 1997. A simple adaptive procedure leading to correlated equilibrium. Game Theory and Information .

Huangfu, Q., Hall, J.J., 2018. Parallelizing the dual revised simplex method. Mathematical Programming Computation 10, 119–142.

Koutsoupias, E., Papadimitriou, C., 1999. Worst-case equilibria, in: Stacs, Springer. pp. 404–413.

Lam, S.K., Pitrou, A., Seibert, S., 2015. Numba: A llvm-based python jit compiler, in: Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC, pp. 1–6.

Lanctot, M., Zambaldi, V., Gruslys, A., Lazaridou, A., Tuyls, K., Perolat, J., Silver, D., Graepel, T., 2017. A unified Game-Theoretic approach to multiagent reinforcement learning. Adv. Neural Inf. Process. Syst. 30.

Lotker, Z., Patt-Shamir, B., Tuttle, M.R., 2008. A game of timing and visibility. Games and Economic Behavior 62, 643–660.

Myerson, R.B., 1997. Game theory: analysis of conflict. Harvard university press.

Nash, J.F., 1950. Equilibrium points in N-Person games. Proc. Natl. Acad. Sci. U. S. A. 36, 48–49.

Neller, T.W., Lanctot, M., 2013. An introduction to counterfactual regret minimization, in: Proceedings of Model AI Assignments, The Fourth Symposium on Educational Advances in Artificial Intelligence (EAAI-2013), pp. 1602–1603.

Smith, J.M., 1974. The theory of games and the evolution of animal conflicts. Journal of theoretical biology 47, 209–221.

Sobol', I., 1990. Quasi-monte carlo methods. Progress in Nuclear Energy 24, 55–61.

Sobol', I.M., 1967. On the distribution of points in a cube and the approximate evaluation of integrals. Zhurnal Vychislitelnoi Matematiki i Matematicheskoi Fiziki 7, 784–802.

Appendix A. Proof of Theorem 2.5

We follow broadly the same scheme as Theorem 3.6 of Lotker et al. (2008), mutatis mutandis since the equations are different. Their proof proceeds by supposing the existence of a Nash equilibrium and deriving properties to characterize it. Contrary to what they claim, we can prove the existence of a Nash equilibrium without any computation using Dasgupta and Maskin (1986). Let (f_1, f_2) be the density functions of Player 1 and 2 in a Nash equilibrium. We note S_1, S_2 their support.

Proposition A.1. *For almost all $x \in S_1$, $f_2(x) \sim \frac{1}{(1-x)^3}$ and conversely.*

Proof. Let us simplify the notations by noting $a \sim f_1$ and $b \sim f_2$ the random moves of players 1 and 2. We note u_1^* the utility of Player 1 and $\bar{b}$ the expectation of b. Then, by Theorem 2.4, for almost all $a \in S_1$,

$$\begin{aligned} u_1^* = u_1(a) &= \mathbb{E}_b[u_1(a,b)] \\ &= \mathbb{E}_b\left[Rb(1-a) - aP + [a > b](1-a)(1-b)R\right] \\ &= R\bar{b}(1-a) - aP + (1-a)\int_0^a (1-b)Rf_2(b)db \end{aligned}$$

$$\int_0^a (1-b)Rf_2(b)db = \frac{u_1^* + aP}{1-a} + R\bar{b}$$

We derivate to obtain $f_2(a) = \frac{u_1^*+P}{R(1-a)^3}$. □

Proposition A.2. *With the exception of a set measure zero, $S_1 = S_2$.*

Proof. We apply the previous result that implies $x \in S_1 \implies f_2(x) \neq 0$. □

Proposition A.3. $\inf S_1 = \inf S_2 = 0$

Proof. Suppose $\inf S_1 = \inf S_2 = l > 0$. By Theorem 2.4, $u_1(l) = u_1^* = \max_x u_1(x)$. But $u_1(0) > u_1(l)$ since playing 0 decreases the risk of Player 1 without compromising its chances to get the reward. By contradiction, $\inf S_1 = \inf S_2 = 0$. □

Proposition A.4. *For all intervals $[x_1, x_2]$ with $0 < x_1 < x_2 < \sup S_1$ we have that $\int_{x_1}^{x_2} f_1(x)dx > 0$*

Proof. Suppose that there is an interval $[x_1, x_2]$ such that $\int_{x_1}^{x_2} f_1(x)dx = 0$. Assume that this interval is maximal so that $x_1, x_2 \in S_1$. We also have that $\int_{x_1}^{x_2} f_2(x)dx = 0$ since $S_1 = S_2$. Hence Player 2 never plays between x_1 and x_2 and this implies that $u_1(x_1) > u_1(x_2)$ since Player 1 can decrease their risk without compromising its chances to get the reward. This contradicts the fact that $u_1(x_1) = u_1(x_2)$ since they are both in the support of the Nash equilibrium. □

Proposition A.5. *There is no point x with positive probability.*

Proof. Suppose the existence of a point x with positive probability (which means that $f_1(x)$ is a Dirac. Since we determined the expression of $f_1(x)$ for almost every x, there is a $\varepsilon > 0$ such that there is no other point with positive probability in $[x, x+\varepsilon]$. Hence, there is $0 < \varepsilon' < \varepsilon$ such that $u_2(x+\varepsilon') > u_2(x)$ since any positive ε' improves the probability of winning the reward by $\mathbb{P}[\text{Player 1 plays } x]$ and the risk increment goes to 0 with ε'. □

Theorem 2.5. *Up to a set of measure zero, the CoR game admits a unique Nash equilibrium. This equilibrium is symmetric and its distribution is* $f(x) = \left[x < 1 - \sqrt{\frac{k-1}{k+1}}\right] \frac{k-1}{(1-x)^3}$ *with* $C := \frac{P}{R}$ *and* $k := \sqrt{(C+1)^2 + 1}$*. The the average move is* $\bar{r} = k - (C+1)$ *and the utility of each player is* $u^* = R\bar{r}$*.*

Proof. Up to a set of measure 0 and on which the probability is 0, we have determined the expression of f_1 and f_2 above up to a constant. This means that $f_1 = f_2 = f$. We still have to find the constant and the upper limit of the support. Let us reiterate the calculations, this time using the fact that $0 \in S$ with S the support of f. This means that for any $a \in S$:

$$\begin{aligned} 0 &= u_1(a) - u_1(0) \\ &= \mathbb{E}_b[u_1(a,b) - u_1(0,b)] \\ &= \mathbb{E}_b\left[Rb(1-a) - aP + [a > b](1-a)(1-b)R - Rb\right] \\ &= -aR\bar{b} - aP + (1-a)\int_0^a R(1-b)f(b)db \end{aligned}$$

$$\int_0^a R(1-b)f(b)db = (R\bar{b} + P)\frac{a}{1-a}$$

We derivate to obtain:

$$f(a) = \left(\bar{b} + \frac{P}{R}\right)\frac{1}{(1-a)^3}$$

We now have two unknowns and two unknowns:

$$\begin{cases} h := \sup S \text{ such that } \int_0^h f(x)dx = 1 \\ \bar{b} = \int_0^h xf(x)dx \end{cases}$$

We define $C := \frac{P}{R}$.

$$\int_0^h f(x)dx = \frac{\bar{b} + C}{2}\left(\frac{1}{(h-1)^2} - 1\right) = 1$$

$$\int_0^h xf(x)dx = \bar{b} = \frac{\bar{b} + C}{2}\frac{h^2}{(h-1)^2}$$

We get

$$\bar{b} = \frac{h^2}{1-(h-1)^2}$$

$$2\frac{(h-1)^2}{1-(h-1)^2} - C = \frac{h^2}{1-(h-1)^2}$$

$$2(h-1)^2 - C + C(h-1)^2 - h^2 = 0$$

$$(C+1)h^2 - (2C+4)h + 2 = 0$$

Hence, $h = \frac{2+C\pm\sqrt{C^2+2C+2}}{1+C}$
Since $h < 1$, we get

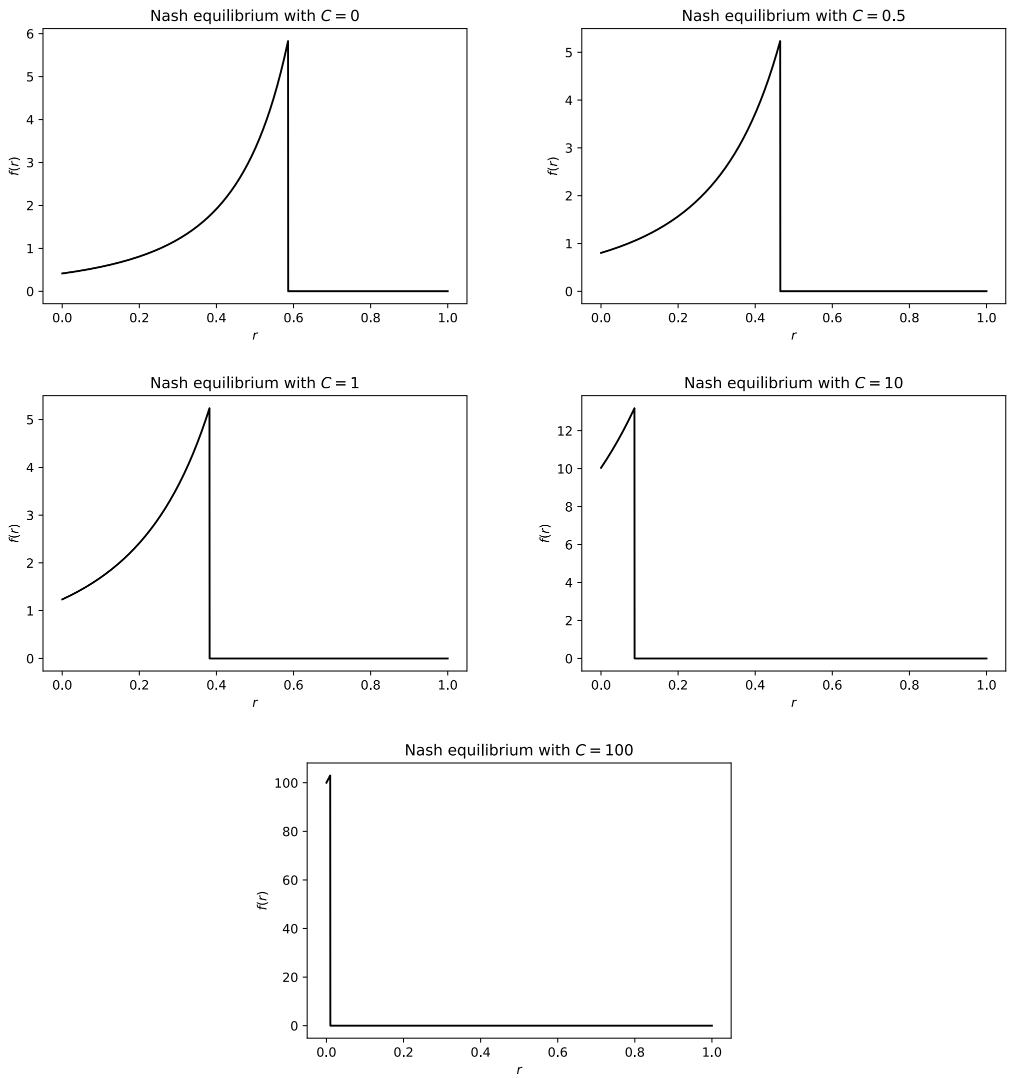
$$\begin{aligned} h &= \frac{2+C-\sqrt{C^2+2C+2}}{1+C} \\ &= 1 - \frac{\sqrt{(C+1)^2+1}-1}{1+C} \end{aligned}$$

We remark that $(\sqrt{(C+1)^2+1}+1)(\sqrt{(C+1)^2+1}-1) = (C+1)^2$
Hence $h = 1 - \sqrt{\frac{\sqrt{(C+1)^2+1}-1}{\sqrt{(C+1)^2+1}+1}}$

$$\begin{aligned} \bar{b} + C &= \frac{2}{\frac{1}{(h-1)^2} - 1} \\ &= \frac{2}{\frac{\sqrt{(C+1)^2+1}+1}{\sqrt{(C+1)^2+1}-1} - 1} \\ &= \sqrt{(C+1)^2+1} - 1 \end{aligned}$$

Finally, $f(x) = \left[x < 1 - \sqrt{\frac{k-1}{k+1}}\right] \frac{k-1}{(1-x)^3}$ with $k := \sqrt{(C+1)^2+1}$
From the expressions above, we get $\bar{b} = k - 1 - C$ and $u^* = R\bar{b}$. $\square$

Appendix B. Plots of Nash equilibrium for various values of C



Appendix C. Estimation of r_{max} and w in the multiplayer setting

We recall the equations:

$$f(x) = \frac{C + w}{(n-1)(1-x)^{2+\frac{1}{n-1}}(Cx+w)^{1-\frac{1}{n-1}}}$$

$$\int_0^{r_{max}} f(x)dx = 1$$
$$\int_0^{r_{max}} xf(x)dx = \bar{r}$$

with

$$w := \bar{r}^{n-1}$$

We could estimate the integrals by numerical integration. However $\bar{r}$ becomes smaller when increasing n, which causes f to have a peak at 0 and renders the integral estimates unreliable. Computations [3] give us

$$\int f(x) = \frac{w + nC(1-x) + Cx}{n(1-x)(C+w)} \sqrt[n-1]{\frac{Cx + w}{1-x}}$$

$$\int xf(x) = \frac{w - nw(1-x) + Cx}{n(1-x)(C+w)} \sqrt[n-1]{\frac{Cx + w}{1-x}}$$

Hence

$$\frac{w + nC(1-r_{max}) + Cr_{max}}{n(1-r_{max})(C+w)} \sqrt[n-1]{\frac{Cr_{max} + w}{1-r_{max}}} - \frac{w + nC}{n(C+w)} \sqrt[n-1]{w} = 1 \tag{C.1}$$

$$\frac{w - nw(1-r_{max}) + Cr_{max}}{n(1-r_{max})(C+w)} \sqrt[n-1]{\frac{Cr_{max} + w}{1-r_{max}}} - \frac{w - nw}{n(C+w)} \sqrt[n-1]{w} = \sqrt[n-1]{w} \tag{C.2}$$

The value of integrals are increasing in r_{max} since $f(x)$ is positive, hence for any value of w we can find the corresponding value of r_{max} by binary search using (C.1). Then we are left with finding the value of w using (C.2) . We observe experimentally that the function $w \to \int_0^{r_{max}(w)} xf(x)dx - \sqrt[n-1]{w}$ seems to have only one root and that it is positive before that root and negative after that root. We therefore use binary search.

Equation (C.2) also defines w as a fixed point and the iterative algorithm $w \leftarrow \left(\int_0^{r_{max}(w)} xf(x)\right)^{n-1}$ also converges, although it seems to sometimes loop between a few close values because of numerical errors.

[3]We used Wolfram Alpha with `ReplaceAll[Integrate[(1/(1 - x)) D[((w + x C)/(1 - x))^(1/(n - 1)), x], x], n -> 42]` for various values of n, found a pattern and checked that the derivatives match.

We remark that w goes very quickly to 0 as n increases. This causes numeric errors in the computations of the primitives at 0 that we fix by storing $\log w$ instead of w. For the same reason, we compute the integrals in log space.

Equation (C.2) is more interesting as the indefinite integral $F(r_{max})$ can be negative. Depending on the sign, we store either $\log F(r_{max})$ or $\log -F(r_{max})$.

Finally, we apply the Log-Sum-Exp Trick to compute the value of the integral in log space. The final algorithm to find w and r_{max} takes less than 1 ms on a laptop for any value of the parameters.

Appendix D. Proof of Theorem 4.3

Theorem 4.3. *For $\rho = 1$, the equilibrium is given by:*

$$p(x) = \frac{1+C}{1-x}\left[x < 1 - \exp\left(-\frac{1}{C+1}\right)\right]$$

We have

$$\bar{r} = 1 - (C+1)\left(1 - \exp\left(-\frac{1}{C+1}\right)\right)$$

For $\rho = -1$, the equilibrium is given by:

$$p(x) = \frac{C}{(1-2x)^{3/2}}\left[x < \frac{1}{2} - \frac{C^2}{2(C+1)^2}\right]$$

We have

$$\bar{r} = \frac{1}{2C+2}$$

Proof. We reuse the proof of Appendix A, only modifying the calculations. We note a and b the actions of the players, aka their individual probability of failure noted r_1 and r_2 above and note c their joint probability of failure noted $\tilde{r}$ above. c is a function $c(a, b, \rho)$. We note p the probability distribution corresponding to the Nash equilibrium. In both cases, we have the equality

$$\int_0^1 p(b)c \; db + aC = \int_0^a p(b)(1 - a - b + c) \; db$$

When $\rho = 1$, $c = \min(a, b)$

$$\int_0^a p(b)b\ db + a\int_a^1 p(b)\ db + aC = \int_0^a p(b)(1-a)\ db$$

We derivate wrt a:

$$ap(a) + \int_a^1 p(b)\ db - ap(a) + C = p(a)(1-a) - \int_0^a p(b)\ db$$

$$p(x) = \frac{1+C}{1-x}$$

We solve $\int_0^{r_{max}} p(x)\ dx = 1$ as $\log(1-r_{max}) = -\frac{1}{C+1}$, thus

$$r_{max} = 1 - \exp\left(-\frac{1}{C+1}\right)$$

Finally, a simple computation gives

$$\bar{r} = 1 - (C+1)\left(1 - \exp\left(-\frac{1}{C+1}\right)\right)$$

When $\rho = -1$, $c = \max(0, a+b-1)$. In other terms, $c = 0$ if $b < 1-a$ and $c = a+b-1$ if $b > 1-a$.

We first treat $a > \frac{1}{2}$, aka $a > 1-a$, to show by contradiction that $p(a) = 0$.

$$\int_{1-a}^1 p(b)(a+b-1)\ db + aC = \int_0^{1-a} p(b)(1-a-b)\ db$$

$$aC = \int_0^1 p(b)(1-a-b)\ db = 1 - \bar{r} - a$$

This is impossible, thus $p(a) = 0$ for $a > 1/2$.

We now suppose $a < 1-a$:

$$\int_{1-a}^1 p(b)(a+b-1)\ db + aC = \int_0^a p(b)(1-a-b)\ db$$

$\int_{1-a}^1 p(b)(a+b-1)\ db = 0$ since $1-a > \frac{1}{2}$. We derivate twice to get:

$$(1-2a)p(a) = \int_0^a p(b)\ db + C$$

$$1 - 2a)p'(a) - 3p(a) = 0$$

The first equation gives $p(0) = C$. Combined with the second differential equation, we get:

$$p(x) = \frac{C}{(1-2x)^{3/2}}$$

Similarly, we solve

$$r_{max} = \frac{1}{2} - \frac{C^2}{2(C+1)^2}$$

$$\bar{r} = \frac{1}{2C+2}$$

□

Appendix E. Proof of Theorem 5.1

Theorem 5.1. *Given a polytope P defined by constraints $c_1, \ldots, c_m$*

$$\Pr\left[\textsc{SumDiamSquared}(K, c_1, \ldots, c_m) < \varepsilon\right] \leq F_{\chi^2}\left(\frac{\varepsilon}{diam(P)}, K\right)$$

with $F_{\chi^2}(\cdot, K)$ the cumulative distribution function of the χ^2 distribution with K degrees of freedom.

Proof. We suppose that the diameter d is obtained in some unit direction $\vec{a}$:

$$\max_{x \in P} a \cdot x - \min_{x \in P} a \cdot x = d$$

We define D the line $\left[\arg\min_{x \in P} a \cdot x, \arg\max_{x \in P} a \cdot x\right]$. Then, for any v,

$$\max_{x \in P} v \cdot x - \min_{x \in P} v \cdot x \geq \max_{x \in D} v \cdot x - \min_{x \in D} v \cdot x = d \times |v \cdot a|$$

Hence,

$$\Pr\left[\max_{x \in P} v \cdot x - \min_{x \in P} v \cdot x < \varepsilon\right] \leq \Pr\left[|v \cdot a| < \frac{\varepsilon}{d}\right]$$

If v is a vector of iid standard Gaussian variables, $v \cdot a \sim \mathcal{N}(0, 1)$ and $(v \cdot a)^2$ follows a χ^2 distribution with 1 degree of freedom. The sum of multiple iterations will follow:

$$\Pr\left[\sum_{i}^{K}(\max_{x\in P} v_i \cdot x - \min_{x\in P} v_i \cdot x)^2 < \varepsilon\right] \leq \Pr\left[\sum_{i}^{K} z_i^2 < \frac{\varepsilon}{d}\right]$$

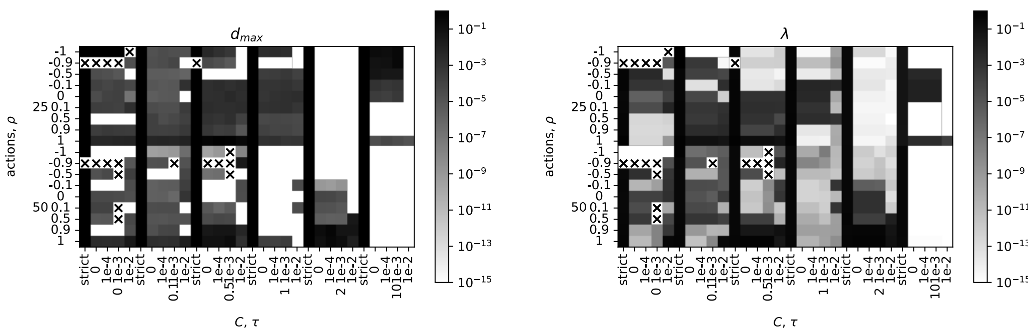
where $z_i \sim \mathcal{N}(0,1)$ are independent normal variables.

$\sum_i^K z_i^2$ follows a χ^2 distribution with K degrees of freedom. With $F_{\chi^2}(x, K)$ the cumulative distribution function, we get:

$$\Pr\left[\sum_{i}^{K}(\max_{x\in P} v_i \cdot x - \min_{x\in P} v_i \cdot x)^2 < \varepsilon\right] \leq F_{\chi^2}\left(\frac{\varepsilon}{d}, K\right)$$

□

Appendix F. Linear equilibrium in finite approximations with `shift = false`



Appendix G. Description of Regret Matching

Regret matching was introduced by Hart and Mas-Colell (1997) as a strategy such that if each player applies it, the observed sequence of plays can converge to a correlated equilibrium. The idea behind regret matching is to adjust the probabilities of playing each action based on the accumulated regrets for each action. We describe the version that was originally named Unconditional Regret Matching.

Regret is defined as the difference between the payoff received from a particular action and the actual historical payoff. The intuition behind regret is that a player may regret not having played a different action that would have resulted in a better payoff. In regret matching, a player starts by assigning equal probability to each possible action. After playing the game and receiving a payoff, the player computes the regret for each action based on the difference between the payoff received and the maximum payoff that could have been obtained by playing a different action. The player then updates the probabilities of each action proportional to their positive regrets. More formally, let $s \in \Sigma_i$ be an action, $r(s)$ be the regret for playing action s, and $\Pr_t(s)$ be the probability of playing action s at time t. The regret-matching update rule is as follows:

$$\Pr_t(s) = \begin{cases} \frac{r^+(s)}{\sum_{s'} r^+(s')} & \text{if } \sum_{s'} r^+(s') > 0 \\ \frac{1}{|\Sigma_i|} & \text{else} \end{cases}$$

where $r^+(s) = \max(r(s), 0)$ and $r(s)$ is the cumulative regret for playing action s up to time t, defined as:

$$r(s) = \sum_{k=1}^{t} u_i(s, s_{-i}[k]) - u_i(s[k])$$

where $s[t]$ is the history of what was played at time t. The intuition behind the regret-matching update rule is that actions that have positive regrets are given higher probabilities in the next round, while actions with zero or negative regrets are given zero probability. Over time, as the player continues to play the game and accumulate regrets, the probabilities of playing each action converge to a correlated equilibrium of the game, which is a set of probabilities over actions that maximizes the player's expected payoff, given the strategies of the other players.

A variant named Counterfactual Regret Minimization described by Neller and Lanctot (2013) does not actually sample actions and instead maintains a cumulative profile that can be viewed as the expectation of moves under the Regret Matching strategy. This deterministic variant converges faster in practice.

TODO: describe equations of CFR, explain tabular computations