Incorporating Ignorance within Game Theory: An Imprecise Probability Approach

Bernard Fares* and Mimi Zhang

School of Computer Science and Statistics, Trinity College Dublin, Ireland

Abstract

Ignorance within non-cooperative games, reflected as a player's uncertain preferences towards a game's outcome, is examined from a Bayesian point of view. This topic has had scarce treatment in the literature, which emphasises exogenous uncertainties caused by other players or nature and not by players themselves. That is primarily because a player's endogenous uncertainty over an outcome poses significant challenges and complex sequences of reciprocal expectations. Therefore, it is often ignored, and preferences are either assumed from a continuous domain or set using introspection, resulting in non-optimal models. We here explore a solution concept based on recent research in imprecise probabilities and de Finetti's approach to defining subjective probabilities, which utilises bets to assess beliefs. The resulting model allows players to be ignorant about their initial preferences and learn about them in repeated games. Furthermore, it permits improving the value of information in these situations. This model is proposed as a possible solution to the problem of utility inference in game-theoretic settings that include uncertainty over outcomes. We demonstrate it through motivating repeated-game problems modified to have uncertainty and through a simulation over a case of extreme ignorance.

Keywords: Game theory; Imprecise probability; Uncertain utility; Nonparametrics

1 Introduction

One of the central and main requirements of decision theories is knowing a decision-maker's preferences. Those preferences are reflected and quantified as *utilities*. As Von Neumann and Morgenstern (1944) outlined, utilities are numerical measurable quantities that are expected to be complete. Furthermore, based on economics by Pareto and Bonnet (1963), these utilities must be comparable. The decision-maker is expected to have a clear intuition of preference between two objects, events, or even a combination of events.

*Corresponding author. Tel.: +353-834747470. E-mail: FaresB@tcd.ie.

Game theory studies the decision-making process in a strategic setting, a situation where two or more decision-makers, known as players, compete or collaborate to maximise their utilities. Game theory relies on pre-determined utilities to decide a player's optimal move. As stated by Luce and Raïffa (1957), a player is required to know their utility towards each possible outcome in the game and the utilities of all other players. Therefore, a player is expected to have explicit knowledge about each outcome, including outcomes that have never been previously experienced. This knowledge is reflected by having a utility function with the game's complete set of outcomes as a domain.

Usually, the utility function is considered to have a fixed form throughout the entire gameplay. Furthermore, it can be either a defined function, *e.g.* an exponential utility, or utility values attributed to each possible outcome. Whichever way, the possibility that a player may not know an outcome's payoff in advance is ignored and either an assumed utility model from a continuous domain is used or utility values are set using introspection. In many cases, these models are not optimal. For instance, in the *temporal sure preference* principle, Chiara et al. (2013) show that prior preferences represent a minimum coherence requirement to link beliefs at different time points. It suggests that if *'you have a sure preference for A over B at (future) time t. Then you should not have a strict preference for B over A now'*; that is, prior belief should be coherent with posterior preferences. Therefore, an assumed utility makes assessing conditional belief towards future preferences very limited, which implies that classical game-theoretic approaches fail to handle situations of ignorance. Moreover, since an assumed utility value remains fixed throughout the entire gameplay, these approaches don't allow the player to be surprised or learn about their preferences once the outcome is experienced.

Given game theory's many applications in economy, finance, biology etc., a state of endogenous uncertainty over outcomes could arise in many cases, *e.g.* when two companies invest in a new market. It is common for different parties to freely enter a game without having experienced its outcomes before, sometimes as a trial attempt to learn about their preferences towards these outcomes. Therefore, stretching the classical game theory to allow for cases of ignorance over an outcome will expand its application. Instead of assuming a known utility, the uncertain outcome should be assessed and a player's prior unknown preferences estimated. Afterwards, these estimates should be dynamically updated when the outcome is experienced.

Models, such as multiple prior (Gilboa and Schmeidler, 1989) or Choquet expected utility (Schmeidler, 1989), were developed to solve the problem of uncertainty within non-strategic settings. However, as per Gajdos et al. (2004), these models have certain limitations. For instance, they assume extreme pessimism, *e.g.* applying the maximin criterion to the initial set of information (Gilboa and Schmeidler, 1989). Furthermore, they don't allow the decision-maker to incorporate available information before making a decision. In practice, representing decision-maker beliefs is essential, especially since Ellsberg (1961) proved that decision-makers prefer better information settings. Flexible approaches have been developed to overcome these limitations, *e.g.* Gajdos et al. (2004) or Troffaes (2007). Nevertheless, these approaches were never formally extended to cover game theory.

Bayesian theories have emerged to account for cases where a player is in a state of incomplete information about other players' utilities in the game. Nevertheless, these theories don't consider a player's incomplete information about their own utilities. This case raises different challenges and complex sequences of reciprocal expectations (Harsanyi, 1967). For instance, in Nau's (1992) operational method to achieve joint coherence and common knowledge of subjective parameters in non-cooperative games that include incomplete and observable information, a dual characterisation of joint rationality is introduced by generalising Harsanyi's (1967) *Bayesian* equilibrium concept. However, in a *Bayesian* equilibrium, each player is assumed to know their actual type, considered a summary of their actions and payoffs, and only uncertain about other players' types (Harsanyi, 1967, p.1811). This uncertainty towards the game's structure is modelled through a common prior distribution over these players' types. Nau (1992) stretches this assumption further by defining belief-revealing monetary payoff functions, *i.e.* gambles, over a set of outcomes composed of players' joint strategies and the states that represent exogenous uncertainties caused by other players or nature and not by players themselves.

Although Nau's operational method is limited to exogenous uncertainties, it includes some limitations that should be considered when modelling any type of uncertainty. First, it shows that an increased number of uncertain states results in a complex model. That is primarily because a game with incomplete information is converted into several, each representing a possible state whose lower probability is set through a *belief* gamble. The complexity could increase further if these states require updating due to acquired information, *e.g.* in repeated games. Second, it doesn't allow simultaneously including lower and upper bounds of a possible state and computing a solution set given this range; only a scalar value within these bounds is permitted. Finally, it doesn't incorporate a mechanism for players to update or learn about an uncertain domain through experience or acquired information.

Computational models have been developed to treat the problem of players' endogenous uncertainty over outcomes. For example, Astanin and Zhukovskaja (2015) and Chakeri et al. (2008) apply *fuzzy logic* to game theory and deal with ignorance using the notion of fuzzy games. Although these models prove to be successful in certain situations, there is still no actual Bayesian model that could help with statistical inference and the quantification of endogenous uncertainty under strategic settings. For example, how confident can a player be that an outcome will generate a particular payoff? The interest here is using Bayesianism to construct a scalable model that allows computing a game's rational solutions based on prior lower and upper previsions of uncertain domains and updating them through experience to posterior previsions. This model would enable meaningful approximations if computation is not feasible and permit the dynamic incorporation of acquired information. Furthermore, it would provide a way to express expert knowledge, which is challenging to reflect in computational models (Chiandotto, 2014). Moreover, interest also lies in examining the consistency of game theory's axiomatic rules under endogenous uncertainty. Von Neumann and Morgenstern's (1944) normative foundations of this theory are not extended to handle this situation. The proposed axioms of behaviour were always studied with the assumption that players know the payoff of each outcome in the game.

2 Preliminaries

2.1 Imprecise Probabilities

Imprecise probabilities is a well-established framework aimed toward quantification and inference under uncertainty; a state of incomplete or imprecise information. Influenced by de Finetti's (1974, 1937) work on subjective probability, Williams (1974, 1975) worked on an early detailed study of the theory. Walley (1991) then developed it into a more mature one.

An attractive theory under imprecise probabilities is that of lower and upper previsions, represented by $\underline{P}(\cdot)$ and $\overline{P}(\cdot)$. They are, respectively, the supremum acceptable buying price and the infimum acceptable selling price of a gamble. Let $\mathcal{X}$ denote a finite set of an experiment's exhaustive and mutually exclusive outcomes. Let $f(\cdot)$ denote an arbitrary reward function defined on $\mathcal{X}$: if x is the outcome of the experiment, then the reward is $f(x)$, denominated in units of a linear utility scale. As the experiment's outcome is random, the reward of the experiment is random. Hence, the reward of the experiment can be interpreted as a gamble. The lower and upper previsions of f are, respectively, $\underline{P}(f)$ and $\overline{P}(f)$. Furthermore, they are considered a subject's lower and upper expectations of f. Whenever they coincide, such that $P(f) := \underline{P}(f) = \overline{P}(f)$, the resulting $P(f)$ is called a linear prevision. It is seen by de Finetti (1975) as the fair price of f. A special case of upper and lower previsions is when the reward function f takes the form of an indicator function, *i.e.* a 0-1 binary vector. Consequently, these previsions become upper and lower probabilities of the event that the outcome of the experiment belongs to $\mathcal{X}$. As $\underline{P}^k(f) = -\overline{P}^k(-f)$ (Walley, 1991), we are going to limit our discussion to lower previsions and use upper previsions whenever deemed necessary.

Let $\mathcal{A}$ denote a subset of $\mathcal{X}$. A conditional lower prevision $\underline{P}(f|\mathcal{A})$ is the supremum buying price of gamble f given $\mathcal{A}$. This conditional lower prevision has two different interpretations. $\underline{P}(f|\mathcal{A})$ could be considered as the supremum buying price of a gamble f whose value is zero outside $\mathcal{A}$. Or, as Walley's (1991) *updating principle* suggests, it could be considered as the updated supremum buying price of f after receiving information that the outcome belonged to $\mathcal{A}$.

A gamble is said to be *desirable* when a subject is willing to accept it whenever offered. This doesn't necessarily suggest that a non-desirable gamble is rejected. The latter only means that the subject is undecided about whether to accept it. Formally, Walley (1991) considers a gamble f to be desirable if $\inf_{x\in\mathcal{X}} f(x) > 0$, *i.e.* when it increases the subject's utility no matter what the outcome is. A gamble f is said to be *almost desirable* if $\forall\ \epsilon > 0$, $f + \epsilon$ is desirable. The set of almost-desirable gambles includes all desirable gambles. Furthermore, as per Walley (1991, Theorem 3.8.4, p. 158), these almost-desirable gambles correspond to some linear previsions that are greater than or equal to zero: if f is almost desirable, then $P(f) \geq 0$.

A lower prevision should not allow any opportunities for riskless profits through smart combinations. It is a rationality requirement known as *coherence*. On a linear space, Walley (1991) characterises it using the following axioms:

- $\underline{P}(f) \geq \inf_{x\in\mathcal{X}} f(x)$;
- $\underline{P}(\lambda f) = \lambda \underline{P}(f)$;
- $\underline{P}(f+g) \geq \underline{P}(f) + \underline{P}(g)$.

Coherent lower previsions are challenging to achieve, especially in a non-linear space. Therefore, the lower and upper previsions theory allows constructing coherent models from assessments that only avoid sure loss. *Avoiding sure loss* is a weaker rationality condition that is easier to satisfy. It can be ascertained using Equation (1), which guarantees at least one outcome to generate a payoff that is greater or equal to zero.

$$\sup_{x\in\mathcal{X}} \sum_{f_i\in F} [f_i(x) - \underline{P}(f_i)] \geq 0, \tag{1}$$

where F is a set of gambles.

Lower previsions that avoid a sure loss can be assessed for coherence using Equation (2), which is the general definition of coherence suggested by Walley (1991).

$$\sup_{x\in\mathcal{X}} \left\{ \sum_{f_i\in F} [f_i(x) - \underline{P}(f_i)] - l_0[f_0(x) - \underline{P}(f_0)] \right\} \geq 0, \tag{2}$$

where l_0 is any positive integer and f_0 is a gamble assessed for coherence.

To imply a new gamble's coherent lower prevision from existing assessments, *natural extension* is used. Let $\underline{E}$ denote the natural extension of $\underline{P}$ on F such that for any gamble f, $\underline{E}(f)$ is its supremum buying price implied from $\underline{P}(f_i)$ through linear operations. Furthermore, $\underline{E}$ dominates $\underline{P}$ on F, which allows correcting any incoherent assessment. For instance, any previously assessed $\underline{P}(f)$ that is strictly lower than the implied lower prevision $\underline{E}(f)$ is deemed incoherent. Formally, natural extension is defined as follows:

$$\underline{E}(f) = \sup \left\{ \omega : f(x) - \omega \geq \sum_{f_i\in F} \lambda_i [f_i(x) - \underline{P}(f_i)] \text{ for some } \omega \in \mathbb{R}, \text{and } \lambda_i \geq 0 \right\}. \tag{3}$$

2.2 Revealed-Rules Matrix

Nau (2011) used imprecise probabilities to create a unified theory between subjective probability, game theory, and other equilibrium models used for games and markets. His approach allows converting a non-cooperative game into a matrix that contains the *revealed rules of the game*. These rules allow generating a convex set of probability distributions that represents the game's equilibria. This section discusses Nau (2011)'s theory and how to generate this matrix, which later on is considered a key element in the suggested model.

Consider a non-cooperative game with K risk-neutral players. Let Λ^k denote a finite set of alternatives available to player k, and $\Phi = \Lambda^1 \times ... \times \Lambda^K$ denote the set of all possible outcomes in the game. Let r^k denote a $|\Phi|$-dimensional vector that represents the payoff of player k as a function of these outcomes. Therefore, for outcome $\boldsymbol{\phi} \in \Phi$, player k's payoff is $r^k(\boldsymbol{\phi})$.

Consider e_i^k to be the event in which player k chooses the alternative $a_i^k \in \Lambda^k$ over any other alternatives. Let r_i^k denote the payoff vector available to player k after making this choice, *i.e.* $r_i^k = (r^k|e_i^k)$. For example, in the game paper-rock-scissors, if player one chooses to play 'rock', this results in a payoff vector r_i^k that has only the payout of outcomes $rock - rock, rock - scissors, rock - paper$.

The occurrence of event e_i^k means that player k would trade any payoff vector r_j^k $(j \neq i)$ for r_i^k, conditional on e_i^k. This trade-off can be translated into an unconditional bet that has a *true* payoff vector of $(r_i^k - r_j^k)e_i^k$. Conversely, suppose a player chooses to publicly accept any small bet whose payoff vector is proportional to $(r_i^k - r_j^k)e_i^k$. In that case, they are making their *true* payoff function common knowledge (Aumann, 1976) at the discretion of any observer. Following the same logic, if all players in the game are willing to accept small conditional bets, they make their *true* payoff function public knowledge. As a result, a matrix that exhaustively lists each player's possible *true* payoff functions is built. This matrix is defined as follows:

Definition. A game's revealed-rules(GRR) matrix, denoted by $\boldsymbol{M}$, represents the *true* payoff function of each player, for each possible bet they could accept. Matrix $\boldsymbol{M}$'s columns are indexed by outcomes in Φ and its rows are indexed by $r_i^k - r_j^k$. A GRR matrix contains all *'commonly-knowable information about the rules that is actually used in determining the equilibria of non-cooperative games.'*

Example 1. Consider the classic game 'battle of the sexes', introduced by Luce and Raïffa (1957), in which a couple disagrees about where to go for entertainment. Going together would yield a better utility than going alone. Furthermore, the one going to their preferred place would have a better utility than the other. Table 1 represents the payoff matrix

Table 1: 'Battle of the sexes' - payoff matrix.

	Left	Right
Top	2,1	0,0
Bottom	0,0	1,2

of this game, where $z = 2$, Λ^1={Top(T), Bottom(B)}, Λ^2={Left(L), Right(R)} and $\Phi = \Lambda^1 \times \Lambda^2$={TL, TR, BL, BR}. For example, the payoffs of outcome TL are: $r^1(TL) = 2$ for player one, and $r^2(TL) = 1$ for player two. Table 2 shows the resulting GRR matrix. For instance, the first row represents the case where player one chooses alternative T over B. If player one makes this choice, they are exposed to the bet $r_T^1 - r_B^1$ whose payoff vector is $(2, -1, 0, 0)$.

Table 2: 'Battle of the sexes' - GRR matrix $\boldsymbol{M}$.

	TL	TR	BL	BR
$r^1_T - r^1_B$	2	-1	0	0
$r^1_B - r^1_T$	0	0	-2	1
$r^2_L - r^2_R$	1	0	-2	0
$r^2_R - r^2_L$	0	-1	0	2

When playing a game, each player k will adopt a strategy. As defined by Von Neumann and Morgenstern (1944), a *pure* strategy is a complete plan that helps a player determine the optimal choice from a set of alternatives, under every possible scenario. A *mixed* strategy is a probability distribution over pure strategies; it allows randomising the choice of pure strategies. A Nash (1950) equilibrium is a state where no player can gain a higher payoff by unilaterally deviating from their strategy. If no Nash equilibria exist in pure strategies, then at least one must exist in mixed strategies. In a mixed Nash equilibrium, each player makes choices based on an independent vector of probabilities. Aumann (1974) studied the impact of players correlating their choices instead. He used a randomisation device on outcomes where they may disagree. A *correlated* mixed strategy showed that it could lead to strictly higher expected payoffs than Nash equilibria. Furthermore, it removes the competitive aspect from non-cooperative games and pushes players to cooperate. Aumann (1987) expanded the theory and defined correlated equilibrium as a function that maps a finite probability space to the set of all possible outcomes Φ. He considers a Nash equilibrium to be a particular case of it. Unlike Nash equilibria, one of the interesting aspects of correlated equilibria is the ease of computing them by simply solving a system of linear inequalities.

Nau and McCardle (1990) show that risk-neutral players should have jointly coherent strategies to play the game rationally. This condition is fulfilled if and only if a correlated equilibrium exists. Whereas for risk-averse players, Nau (2011) shows that this rationality condition is fulfilled if and only if a risk-neutral equilibrium exists. Let $\boldsymbol{\rho}$ denote a probability vector in which an element $\rho_{\boldsymbol{\phi}}$ is the probability of the outcome $\boldsymbol{\phi} \in \Phi$. Using the GRR matrix $\boldsymbol{M}$, players are rational if and only if

$$\begin{cases} \boldsymbol{M}\boldsymbol{\rho} \geq 0, \\ \boldsymbol{\rho} \geq 0, \\ \boldsymbol{\rho}' 1 = 1. \end{cases} \tag{4}$$

The system of linear inequalities (4) defines a convex polytope of correlated equilibria. It is the bounded intersection of a finite set of closed half-spaces and is considered to contain the rational solutions of the game.

Example 2. Continuing the 'battle of the sexes' example, the system of linear inequalities

(4) results in the following polytope C.

$$C = \begin{cases} 2\rho_{TL} - \rho_{TR} \geq 0, \\ -2\rho_{BL} + \rho_{BR} \geq 0, \\ \rho_{TL} - 2\rho_{BL} \geq 0, \\ -\rho_{TR} + 2\rho_{BR} \geq 0, \\ \rho_{\boldsymbol{\phi}} \geq 0, \forall\, \boldsymbol{\phi} \in \{\text{TL, TR, BL, BR}\}, \\ \rho_{TL} + \rho_{TR} + \rho_{BL} + \rho_{BR} = 1. \end{cases}$$

As seen in Figure 1, C is a hexahedron with five vertices listed in Table 3.

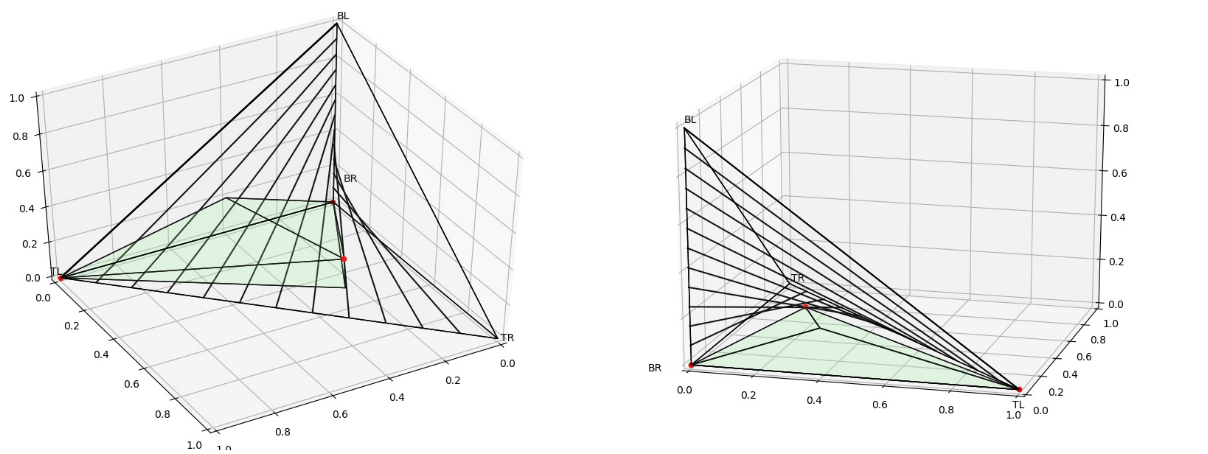

Figure 1: 'Battle of the sexes' - correlated equilibria polytope and the set of all joint probability distributions that are independent between players. Intersections TL and BR are pure Nash equilibria, and intersection (2/3 T, 1/3 B)×(1/3 L, 2/3 R) is a mixed Nash equilibrium.

Table 3: 'Battle of the sexes' - polytope vertices.

	TL	TR	BL	BR
Vertex 1	1	0	0	0
Vertex 2	0	0	0	1
Vertex 3	2/9	4/9	1/9	2/9
Vertex 4	2/5	0	1/5	2/5
Vertex 5	1/4	1/2	0	1/4

Let N denote the number of possible outcomes in Φ. In the system of linear inequalities (4), the last two constraints, $\boldsymbol{\rho} \geq 0$ and $\boldsymbol{\rho}'1 = 1$, define a $N-1$ dimensional simplex, containing all probability distributions on outcomes. The polytope defined by all constraints is a subset of the simplex and contains all correlated equilibria. If the polytope has a dimension smaller than $N-1$, the distribution of correlated equilibria will lie on its boundary. Let

I denote the set of all joint probability distributions of independent variables (here, players). It is the system of nonlinear constraints, $I = \{\boldsymbol{\rho} : \rho_{\boldsymbol{\phi}} = \rho_1(a^1) \times ... \times \rho_z(a^z), \forall \boldsymbol{\phi} \in \Phi\}$, where $a^k \in \Lambda^k$, and ρ_k denotes the marginal probability distribution on Λ^k. As described by Nau et al. (2004), in a 2×2 game, the simplex is a 3-dimensional tetrahedron and I is a 2-dimensional saddle. The set of Nash equilibria is the intersection of I and the correlated equilibria polytope. Nash equilibria only rest on the surface of this polytope. In Example 2, Figure 1 shows three intersections. Intersections TL and BR are pure Nash equilibria, and the intersection (2/3 T, 1/3 B)×(1/3 L, 2/3 R) on the inefficient frontier is a mixed Nash equilibrium.

2.3 Non-parametric Utility Updating

Based on Augustin and Coolen's (2004) work, Houlding and Coolen (2012) introduced a non-parametric predictive utility inference (NPUI) framework for utility induction under extreme ignorance. Their work considers decision-making within non-strategic settings. For sequential decision problems, NPUI features an interesting updating mechanism that assesses the impact of additional observations on utility previsions. This updating mechanism allows the creation of a learning model that helps a decision-maker adjust their utility towards a novel outcome once it is experienced.

The updating mechanism is based on assumption $\mathrm{A}_{(n)}$ proposed by Hill (1968, 1988, 1993). $\mathrm{A}_{(n)}$ is particularly useful for predictions with extremely vague prior knowledge of the underlying distribution. It assumes that pre-observations are exchangeable (de Finetti, 1974). For instance, for two random variables Y_1 and Y_2, $P(Y_1 = y_1, Y_2 = y_2) = P(Y_1 = y_2, Y_2 = y_1)$ holds for all values y_1 and y_2. Furthermore, one of its main pillars is assigning equal mass to the probability that a post-observation falls in $n+1$ distinct intervals created by n observations on a domain $\mathbb{R}$.

In NPUI, observations are restricted to utility values and are bound to the interval $[0, 1]$. This is to avoid having an outcome infinitely better or worse than other alternatives. The values 0 and 1 are, respectively, the worst and best utilities of two actual or 'hypothetical' outcomes. Furthermore, the mechanism assumes that pre-observed utility values are exchangeable under a collection of outcomes. Those outcomes are considered to be sensibly grouped under a particular taxonomic category, *e.g.* sports, computer brands, etc.

Consider a set of ordered known utilities $u_{(1)}, u_{(2)}...u_{(n)}$, such that $0 < u_{(i)} < u_{(i+1)} < 1$. This splits $[0, 1]$ into $n + 1$ intervals. Based on assumption $\mathrm{A}_{(n)}$, a novel outcome whose utility is exchangeable with the existing known utilities has a probability $\frac{1}{n+1}$ of falling in one of these intervals. This is then extended by NPUI to lower and upper utility previsions. Let U_{new} denote a pre-observed utility value. Its lower and upper previsions are defined as follows:

$$\underline{P}(U_{new}) = \frac{1}{n+1} \sum_{i=1}^{n} u_i, \tag{5}$$

$$\overline{P}(U_{new}) = \frac{1}{n+1} \left(1 + \sum_{i=1}^{n} u_i\right) = \frac{1}{n+1} + \underline{P}(U_{new}). \tag{6}$$

Equations (5) and (6) show that the difference between the lower and upper previsions of U_{new} is $\frac{1}{n+1}$. This indicates that when the number of experienced outcomes increases, NPUI reduces the range of possible values that the expected utility of a novel outcome can take. Furthermore, Equations (5) and (6) suggest that in the extreme case where there are no previously experienced outcomes with known and exchangeable utilities, the expected utility can take any value in the range $(0, 1)$. In the opposite case, where the number of experienced outcomes is infinite, lower and upper previsions coincide, which indicates that the expected utility is identified.

If the observed utility of U_{new} is u_{new}, adding it to the set of known utilities will make it fall in one of the existing intervals. Then, the probability of another novel outcome having its pre-observed utility, U^*_{new}, falling in one of the updated intervals is $\frac{1}{n+2}$. Formally, this leads to the following equations.

$$\underline{P}(U^*_{new}|u_{new}) = \frac{1}{n+2}\left(\sum_{i=1}^{n} u_i + u_{new}\right) = \frac{n+1}{n+2}\underline{P}(U_{new}) + \frac{u_{new}}{n+2}, \tag{7}$$

$$\overline{P}(U^*_{new}|u_{new}) = \frac{1}{n+2}\left(1 + \sum_{i=1}^{n} u_i + u_{new}\right) = \frac{n+1}{n+2}\overline{P}(U_{new}) + \frac{u_{new}}{n+2}. \tag{8}$$

Equations (7) and (8) highlight how a novel outcome's lower and upper utility previsions update when this outcome is experienced. This updating proves to be useful in a repeated decision-making situation. It allows the decision-maker to improve their utility's lower and upper previsions towards exchangeable outcomes.

3 Suggested Model Under Uncertainty

Work done by Nau (2011) on extending de Finetti's (1974, 1937) subjective probability approach to non-cooperative games, coupled with the imprecise probabilities toolkit, provides a compelling framework that could serve the problem of ignorance within games. It allows us to convert a game to a matrix that shows the commonly-knowable information required to determine equilibria and subsequently include a player's lower and upper expectations regarding the source of uncertainty, *i.e.* the unknown outcome's payoff. However, this would require an approach to represent the source of uncertainty. Walley (1991) provides a compelling and straightforward methodology to do that. He models uncertainty as a gamble bound to the possibility space of an experiment. The true state of this experiment determines the gamble's uncertain reward. The possibility space should have mutually exclusive states that are detailed enough to describe the subject's beliefs towards the domain of interest. Therefore, a *pragmatic* possibility space can be used to include theoretical as well as observable states. It is not required to be exhaustive and should include sufficiently important and practically possible states, *i.e.* states with a non-zero probability of happening. Since beliefs are often incomplete under uncertainty, this possibility space can evolve and get reformulated to include new pragmatic possibilities. Inspired by Walley (1991), an elicitation model that uses a player's pre-existing beliefs towards the gamble's possible outcomes allows determining its supremum buying and infimum selling prices. Furthermore, Houlding and Coolen's (2012) non-parametric predictive utility inference framework can be used as a dynamic updating mechanism for repeated games. It allows a player to change their preferences towards an unknown outcome based on new observations and hence make more desirable moves.

3.1 Enhanced Revealed-Rules Matrix

Consider a game where players are risk neutral. In practice, the payoff vectors r^k ($k = 1, \ldots, z$) are not always fully known. We here develop a method for constructing a state of common knowledge of the key parameters of the payoff functions. In particular, whenever the payoff $r^k(\cdot)$ for an outcome $\boldsymbol{\phi} \in \Phi$ is unknown, we introduce an experiment with the outcome set $\mathcal{X}$. Let $f(\cdot)$ denote an arbitrary reward function defined on $\mathcal{X}$. Whenever a payoff $r^k(\boldsymbol{\phi})$ is uncertain, it will be replaced by a gamble: $r^k(\boldsymbol{\phi}) = f(x)$, where $x \in \mathcal{X}$ is random and the probability distribution over x is unknown. This allows modelling ignorance using a source of uncertainty that represents the pragmatic possibility space a player can face when landing an outcome with an unknown payoff. If the strategic setting involves more than one source of uncertainty, an experiment is required for each.

It is essential to note that $\mathcal{X}$ is assumed free from any uncertainty. Each state $x \in \mathcal{X}$ becomes a potential consequence if the game results in an uncertain outcome, *i.e.* a gamble. Furthermore, a player must have complete and consistent individual preferences across the domains of available gambles and known outcomes in the game. For example, if a state can arise through gamble f, gamble g or a known outcome, its payoff is expected to be the same in all cases. Although beyond the scope of this paper, Jansen et al. (2022) suggest two user-friendly and robust *preference systems* that could assist a player with assessing their preferences over possible states and outcomes in the game, especially when indecisive. Their method relies on a few ranking questions that allow setting ordinal preferences. Then two different approaches can be used to determine the cardinality of these preferences. The first is a *time elicitation* approach based on the player's consideration time for ranking two states. The second is a *label elicitation* approach that relies on the player to assign pre-defined labels of preference strength.

Example 3. Continuing the 'battle of the sexes' example, assume that the payoff of the outcome TL is unknown to player one; that is, $r^1(\text{TL})$ is unknown. We hence utilize a gamble f to determine $r^1(\text{TL})$: if the outcome of the experiment is x, then $r^1(\text{TL}) = f(x)$. Table 4 represents the resulting payoff matrix of this modified version of 'battle of the sexes'.

Table 4: The modified version of 'battle of the sexes' - payoff matrix.

	L	R
T	$f, 1$	$0, 0$
B	$0, 0$	$1, 2$

By leveraging on the work done by Nau (2011), the game can be transformed into a GRR matrix. However, the existing theory doesn't support cases of uncertainty. Therefore, the model should be extended. Recall, in the event where player k chooses the alternative a_i^k over any other alternative a_j^k, they are practically making a bet that is equivalent to buying the payoff vector r_i^k and selling r_j^k. However, with the presence of ignorance, if the payoff $r_i^k(\boldsymbol{\phi}) \in r_i^k$ or $r_j^k(\boldsymbol{\phi}) \in r_j^k$ is unknown, it is replaced with a gamble. Continuing Example 3, this replacement results in Table 5.

Table 5: The modified version of 'battle of the sexes' - GRR matrix with an uncertain payoff modelled as gamble f.

	TL	TR	BL	BR
$r_T^1 - r_B^1$	f	-1	0	0
$r_B^1 - r_T^1$	0	0	$-f$	1
$r_L^2 - r_R^2$	1	0	-2	0
$r_R^2 - r_L^2$	0	-1	0	2

At this stage, the GRR matrix is incomplete. It requires assessing and valuing the underlying gambles. The value of a gamble is considered to be the supremum or infimum price it is bought or sold for. Let $\underline{P}^k(f)$ and $\overline{P}^k(f)$ be respectively the lower and upper previsions chosen by player k for gamble f. This means a player would be willing to pay $\alpha(\underline{P}^k(f) - \epsilon)$ or get paid $\alpha(\overline{P}^k(f) + \epsilon)$, in exchange for an uncertain reward αf, where $\epsilon \geq 0$ and α is a small positive number. That said, the GRR matrix can be enhanced as follows:

LEMMA 1. In a non-cooperative game where endogenous uncertainty over one or several outcomes exists, an enhanced, more generic form of the revealed-rules matrix is achieved when the payoff $r_i^k(\boldsymbol{\phi})$ of each outcome ϕ in the bought vector r_i^k is replaced with its lower prevision $\underline{P}^k(r_i^k(\boldsymbol{\phi}))$, and the payoff $r_j^k(\boldsymbol{\phi})$ of each outcome ϕ in the sold vector r_j^k is replaced with its upper prevision $\overline{P}^k(r_j^k(\boldsymbol{\phi}))$. Furthermore, the following properties apply:

- $\forall \boldsymbol{\phi} \in \Phi$, if $r_i^k(\boldsymbol{\phi})$ does not represent a gamble, $\underline{P}^k(r_i^k(\boldsymbol{\phi})) = \overline{P}^k(r_i^k(\boldsymbol{\phi})) = r_i^k(\boldsymbol{\phi})$;
- $\forall \boldsymbol{\phi} \in \Phi$, if $r_j^k(\boldsymbol{\phi})$ does not represent a gamble, $\underline{P}^k(r_j^k(\boldsymbol{\phi})) = \overline{P}^k(r_j^k(\boldsymbol{\phi})) = r_j^k(\boldsymbol{\phi})$;
- if $r_i^k(\boldsymbol{\phi})$ is a sold gamble, *i.e.* $r_i^k(\boldsymbol{\phi}) = -f$, then $\underline{P}^k(r_i^k(\boldsymbol{\phi}))$ is equal to $-\overline{P}^k(-r_i^k(\boldsymbol{\phi}))$;
- if $r_j^k(\boldsymbol{\phi})$ is a sold gamble, *i.e.* $r_j^k(\boldsymbol{\phi}) = -f$, then $\overline{P}^k(r_j^k(\boldsymbol{\phi}))$ is equal to $-\underline{P}^k(-r_j^k(\boldsymbol{\phi}))$;

, where f denotes an arbitrary payoff function defined on outcome ϕ's possibility space.

Now, the resulting model permits cases of ignorance. Applying it to the modified version of 'battle of the sexes' returns the GRR matrix in Table 6. The enhanced GRR matrix is interpreted as a system of inequalities that returns all correlated equilibria of a game given the specified lower and upper previsions of the underlying payoffs. This essentially

Table 6: The modified version of 'battle of the sexes' - enhanced GRR matrix, where gamble f is replaced by its lower and upper previsions.

	TL	TR	BL	BR
$r_T^1 - r_B^1$	$\underline{P}^1(f)$	-1	0	0
$r_B^1 - r_T^1$	0	0	-$\overline{P}^1(f)$	1
$r_L^2 - r_R^2$	1	0	-2	0
$r_R^2 - r_L^2$	0	-1	0	2

means that adopting a correlated strategy depends on the player's choice of value amongst the range of possible valuations an uncertain outcome can have. Therefore, a choice rule is required. For instance, Houlding and Coolen (2012) propose two decision-making rules that rely on the decision-maker's level of pessimism. The first is based on the attitude of Extreme Pessimism and requires choosing the outcome or sequential decision path whose lower prevision is greatest. The second is based on the attitude of Extreme Optimism and requires choosing the outcome or sequential decision path whose upper prevision is greatest.

3.2 First Assessment and Refinement

Under extreme ignorance, when imprecision is at its maximum, *vacuous previsions* can be used to value gambles. They are defined as $\underline{P}^k(f) = \inf_{x\in\mathcal{X}} f(x)$ and $\overline{P}^k(f) = \sup_{x\in\mathcal{X}} f(x)$ and proven by Walley (1991) to be coherent, as they respect the coherence requirements listed in Section 2.1. However, using them to model prior beliefs will lead to vacuous posterior previsions. Usually, a player would have some prior information about a gamble, which can be used to increase the accuracy of their previsions. Therefore, amongst several methods provided by the imprecise probabilities toolbox, *general elicitation* (Walley, 1991, p.168) can be used to improve a vacuous assessment. It allows modelling pre-existing beliefs by translating them into explicit judgements. It is by no means a complete method that could cover all practical examples. Nevertheless, it is sufficient enough to build our model.

First, a player starts by making qualitative judgements on elementary events in $\mathcal{X}$. These judgements can be comparative, *e.g.* an event is more probable than the other, or classificatory, *e.g.* an event is probable. Afterwards, judgements are modelled as almost-desirable gambles. For instance, stating that outcome x_1 is probable, means that a player is willing to accept x_1 with odds better than even money. This is equivalent to accepting an almost-desirable gamble $(\boldsymbol{\delta}_{x_1} - \mu)$ with a price $\mu \leq \frac{1}{2}$. Let D^k denote the set of almost-desirable gambles resulting from judgements made by player k. The following is a list of judgement examples and their relevant almost-desirable gambles:

- If outcome x_1 is probable, then gamble $\boldsymbol{\delta}_{x_1} - \frac{1}{2} \in D^k$;
- If outcome x_1 is λ times as probable as outcome x_2, then gamble $\boldsymbol{\delta}_{x_1} - \lambda\boldsymbol{\delta}_{x_2} \in D^k$, where $\lambda \in \mathbb{R}$;
- If outcomes x_1 and x_2 are equally likely, then gamble $\boldsymbol{\delta}_{x_1} - \boldsymbol{\delta}_{x_2} \in D^k$ and $\boldsymbol{\delta}_{x_2} - \boldsymbol{\delta}_{x_1} \in D^k$.

Using this elicitation process, the player should be able to construct any judgement that represents genuine belief and model it as an almost-desirable gamble, denominated in units of probability currency. Once the set of almost-desirable gambles D^k is established, the second stage is to use Equation (9) to check that it avoids a sure loss. This equation takes a more straightforward form than Equation (1).

$$\sup_{x \in \mathcal{X}} \sum_{d_q \in D^k} d_q(x) \geq 0, \tag{9}$$

where $d_q \in D^k$ is an almost-desirable gamble. If D^k is proven to avoid sure loss, the final stage is to compute the relevant lower prevision of each gamble $f \in F$.

Let $K(D^k)$ denote a closed convex set of linear previsions of all gambles $d_q \in D^k$. It is the intersection of all closed convex half spaces determined by the mass function $(P^k(x_1), P^k(x_2), ..., P^k(x_{|\mathcal{X}|}))$ of each $P^k(d_q)$. Since gambles in D^k are almost-desirable, $P^k(d_q) \geq 0$, $\forall d_q \in D^k$. The geometry of $K(D^k)$ is a polytope on the probability simplex, satisfying the set of linear constraints applied to the possibility space $\mathcal{X}$. It is a credal set characterised by having a finite number of extreme points. Using the *lower envelope theorem*, a relationship between linear previsions and coherent lower previsions can be established. The theorem suggests that coherent lower previsions $\underline{P}^k$ of gambles in D^k are none other than the lower envelope of $P^k \in K(D^k)$. This is formally reflected in the following equation:

$$\underline{P}^k(d_q) = \min\{P^k(d_q) : P^k \in K(D^k)\}. \tag{10}$$

Let $\mathcal{E}^k$ denote the natural extension of D^k. It is defined as $\mathcal{E}^k = \sum_{d_q \in D^k} \lambda_q d_q$, where $\lambda_q \geq 0$. $\mathcal{E}^k$ will contain all gambles f whose $P^k(f) \geq 0$. Based on an elementary property of polyhedral cones (Gale, 1960, Theorem 2.13), $K = K(D^k) = K(\mathcal{E}^k)$ is the convex hull of a finite set of linear previsions, and $\mathcal{E}^k$ contains all gambles f whose P^k belongs to this set. Furthermore, the elements of the latter can be considered the extreme points of K. Therefore, it is denoted by $\mathrm{ext}(K)$. Now, lower previsions of gambles in F can be computed by simply taking the lower envelope of $\mathrm{ext}(K)$. This results in Equation (11).

$$\underline{P}^k(f) = \min\{P^k(f) : \forall P^k \in \mathrm{ext}(K)\}. \tag{11}$$

3.3 Dynamic Updating

In a repeated game, whenever a player reaches an outcome that has an uncertain payoff, they get to experience it. In the suggested model, this is reflected by receiving the reward of a gamble. Each time that outcome is experienced, the player develops a preference for it. This preference evolution should be reflected through an update to the lower and upper previsions of the outcome's payoff. Therefore, the lower and upper previsions of the underlying gamble should be updated. Eventually, these previsions will coincide and converge to the fair value of the gamble.

Let n be the total number of observed values of a gamble f, where f represents the uncertain payoff of an outcome $\boldsymbol{\phi}$; that is, $r^k(\boldsymbol{\phi}) = f(x)$ where $x \in \mathcal{X}$ is random and the probability distribution over x is unknown. The lower and upper previsions of $r^k(\boldsymbol{\phi})$ are $\underline{P}^k(f)$ and $\overline{P}^k(f)$. When $n = 0$, these previsions are elicited. However, during each gameplay, whenever $\boldsymbol{\phi}$ is experienced, a new payoff $r^k(\boldsymbol{\phi})$ is observed. This payoff is an arbitrary utility value $u_n = f(x)(x \in \mathcal{X}$ is random). Hence, the player learns more about the outcome's expected payoff. Given such information, the NPUI updating mechanism offers a simple yet robust way to improve the lower and upper previsions of $r^k(\boldsymbol{\phi})$. Nonetheless, applying it in a strategic setting requires some modifications.

First, the NPUI model is initially developed on a unit interval $[0, 1]$, where 0 and 1 are, respectively, a decision-maker's utilities for hypothetical worst and best outcomes. Since a utility function is unique up to a positive linear transformation, the model is applied to the finite interval $[a, b]$, where a and b are respectively the worst and best payoffs of gamble f.

Second, in a strategic setting, the game's outcomes do not necessarily belong to the same collection or can be grouped under the same taxonomic category, *e.g.* sports games. Therefore, we won't adopt an exchangeability assumption across their utility values. It can be argued that observed outcomes whose utility values are exchangeable with that of gamble f might exist outside the game. In this case, we believe such information should be reflected in the gamble's possibility space or the elicitation model discussed in Section 3.2. This allows estimating utility values based on past experiences and permits integrating other pre-existing beliefs towards the novel outcome. Therefore, Equations (5) and (6) are enhanced to include an elicited component, in this case, the one devised in Equation (11). Equations (12) and (13) are the result of this enhancement, which is particularly useful when no observations related to gamble f are available.

LEMMA 2. Let $\{u_1, ..., u_i, ..., u_n\}$ denote a set of known utilities, where n is the total number of observations. In a non-cooperative game, let f denote an arbitrary utility function defined on the possibility space of an uncertain outcome Φ whose utility is exchangeable with the existing known utilities. Let $\underline{P}^k(f|u_0)$ and $\overline{P}^k(f|u_0)$ denote, respectively, player k's initial lower and upper previsions of f when no observations exist, *i.e.* $n = 0$. Based on assumption $\mathrm{A}_{(n)}$, the lower and upper previsions of the pre-observed value of f are as follows:

$$\underline{P}^k(f) = \frac{1}{n+1}\left(\underline{P}^k(f|u_0) + \sum_{i=1}^{n} u_i\right), \tag{12}$$

$$\overline{P}^k(f) = \frac{1}{n+1}\left(\overline{P}^k(f|u_0) + \sum_{i=1}^{n} u_i\right). \tag{13}$$

The elicited components in Lemma 2 are a positive linear transformation of the unit interval $[0, 1]$ of Equations (5) and (6), where 0 and 1 are, respectively, the worst and best utilities available. Therefore, the proof provided by Houlding and Coolen (2012) is still applicable. Equations (12) and (13) show that when $n = 0$, the elicited previsions $\underline{P}^k(f|u_0)$ and $\overline{P}^k(f|u_0)$ are the respective lower and upper previsions. However, once $\boldsymbol{\phi}$ is experienced and an actual payoff is observed, uncertainty regarding its underlying gamble is partially eliminated. In this case, the lower and upper previsions of the subsequent plays are governed by the NPUI framework.

Third, the NPUI model assumes no future outcome is better (worse) or equal to the hypothetical best (worst) outcome. In our approach, a weaker assumption is used. A future outcome can have the same utility as this best (worst) outcome, *i.e.* $a \leq u_i \leq b$.

Finally, we consider that an uncertain outcome in the game should be experienced several times before formulating a proper preference towards it. Hence, the outcome $\boldsymbol{\phi}$ might not necessarily have the same payoff every time it is observed. In practice, such flexibility is required in a strategic setting, especially towards unknown outcomes. In many circumstances, a sole experience does not reflect actual preference. The player should be allowed to try an unfamiliar outcome several times and be surprised about its payoff. This can be achieved as follows:

LEMMA 3. Let $\{u_1, ..., u_i, ..., u_n\}$ denote a set of known utilities, where n is the total number of observations. In a non-cooperative game, let f denote an arbitrary utility function defined on the possibility space of an uncertain outcome Φ whose utility is exchangeable with the existing known utilities. Based on assumption $\mathrm{A}_{(n)}$, when a new exchangeable utility u_{n+1} is observed, player k's lower and upper previsions of the pre-observed value of f, respectively, $\underline{P}^k(f|u_n)$ and $\overline{P}^k(f|u_n)$ are updated as follows:

$$\begin{aligned}\underline{P}^k(f|u_{n+1}) &= \frac{1}{n+2}\left(\underline{P}^k(f|u_0) + \sum_{i=1}^{n} u_i + u_{n+1}\right) \\ &= \frac{n+1}{n+2}\underline{P}^k(f|u_n) + \frac{u_{n+1}}{n+2},\end{aligned} \tag{14}$$

$$\begin{aligned}\overline{P}^k(f|u_{n+1}) &= \frac{1}{n+2}\left(\overline{P}^k(f|u_0) + \sum_{i=1}^{n} u_i + u_{n+1}\right) \\ &= \frac{n+1}{n+2}\overline{P}^k(f|u_n) + \frac{u_{n+1}}{n+2}.\end{aligned} \tag{15}$$

Lemma 3 leverages Equations (7) and (8) by replacing NPUI's lower and upper previsions with the previsions provided by Lemma 2, which include an elicited component. Therefore, the proof provided by Houlding and Coolen (2012) is still applicable. Equations (14) and (15) show that the updated lower and upper previsions are a weighted sum of their respective values before and after observing the payoff u_{n+1}. Such updating seems intuitive. If u_{n+1} falls below the assessed lower prevision, it decreases both lower and upper previsions. If it falls above the upper prevision, it increases both lower and upper previsions. However, if it falls in between, it leads to an increase in the lower prevision and a decrease in the upper prevision. It should be noted that the weights used in Equations (14) and (15) significantly impact how new observations are handled. The increase of existing observations will have a diminishing effect on new ones, which is sensible in a repeated game context. In practice, the early experience of an uncertain outcome greatly influences future game plays.

Updated previsions should comply with the rationality requirements discussed in Section 2.1. According to Augustin and Coolen (2004, Theorem 1), lower and upper previsions based on applying assumption $A_{(n)}$ to observed data are totally monotone, and this total-monotonicity leads to coherence. Hence, Equations (12) and (13) result in coherent lower and upper previsions, assuming that the elicited previsions are also coherent. Furthermore, Augustin and Coolen (2004, Theorem 7) show a strong internal consistency property in the non-parametric updating mechanism, therefore, allowing the coherence argument to be extended to cover Equations (14) and (15).

4 Examples

4.1 First Assessment and Refinement

In the modified version of the game 'battle of the sexes' discussed in Section 3.1, consider that alternatives T and L stand for going to a hockey game, whereas B and R stand for going to the cinema. Furthermore, assume that player one is not familiar with hockey. Hence, their preference towards it is uncertain and replaced by gamble f. The pragmatic possibility space $\mathcal{X}$ can be defined as $\mathcal{X}$={Good(G), Neutral(N), Bad(B)}. $\mathcal{X}$ represents any practically possible state a player could experience by going to the hockey game. The gamble f is assigned the following payoffs f={G:2, N:1, B:0}, where each payoff represents player one's utility over the relevant state. Now, the payoff matrix in Table 4 can be transformed into an enhanced GRR matrix. The latter is reflected in Table 6, where $\underline{P}^1(f)$ and $\overline{P}^1(f)$ are respectively player one's lower and upper previsions of f.

To assess gamble f's lower and upper previsions, player one can rely on previous experiences, which might not be related to hockey, to make qualitative judgements on elementary events in $\mathcal{X}$. Afterwards, these judgements can be converted to almost-desirable gambles under the set D^1. For example:

- Since they generally like sports, having a good experience is probable. This corresponds to a gamble $d_1 = \boldsymbol{\delta}_G - \frac{1}{2} \in D^1$;
- Since they rarely had a bad experience at sports games in the past, having a bad experience is improbable. This corresponds to a gamble $d_2 = \frac{1}{2} - \boldsymbol{\delta}_B \in D^1$;
- Since they usually like sports more than the cinema, a good experience is at least as probable as a neutral one. This corresponds to a gamble $d_3 = \boldsymbol{\delta}_G - \boldsymbol{\delta}_N \in D^1$;
- The odds against a neutral experience are no more than 3 to 1. This corresponds to a gamble $d_4 = \boldsymbol{\delta}_N - \frac{1}{3} \in D^1$.

Applying Equation (9) to the set of gambles $D^1 = \{d_1, d_2, d_3, d_4\}$ shows that it avoids a sure loss. Since the linear prevision of each almost-desirable gamble in D^1 is greater or equal to zero and is determined by its mass function $(P^1(G), P^1(B), P^1(N))$, a credal set $K(D^1)$ can be built. It is the intersection of the following half-spaces.

$$K(D^1) = \begin{cases} P^1(d_1) = P^1(G) - \frac{1}{2} \geq 0 \\ P^1(d_2) = \frac{1}{2} - P^1(B) \geq 0 \\ P^1(d_3) = P^1(G) - P^1(N) \geq 0 \\ P^1(d_4) = P^1(N) - \frac{1}{3} \geq 0 \\ P^1(G), P^1(B), P^1(N) \geq 0 \\ P^1(G) + P^1(B) + P^1(N) = 1 \end{cases}$$

The probability simplex in Figure 2 shows $K(D^1)$ and its corresponding linear previsions on the possibility space $\mathcal{X}$. The equilateral triangle has a height of one, and the probability of each state is identified with perpendicular distances from each side of it. The hyperplane of each gamble in D^1 cuts the simplex into a half-space. The resulting area, coloured in red, is a polyhedron that represents $K(D^1)$. Its intersections are the extreme points, ext$(K(D^1))$ =$\{(\frac{2}{3}, 0, \frac{1}{3}), (\frac{1}{2}, \frac{1}{6}, \frac{1}{3}), (\frac{1}{2}, 0, \frac{1}{2})\}$. The coherent lower and upper probabilities $(\underline{P}^1, \overline{P}^1)$ of each state in $\mathcal{X}$ are the lower and upper envelopes of ext$(K(D^1))$. Hence, $(\frac{1}{2}, \frac{2}{3})$ for 'good', $(0, \frac{1}{6})$ for 'bad', and $(\frac{1}{3}, \frac{1}{2})$ for 'neutral'. The coherent lower prevision of gamble f is computed using Equation (11) as follows:

$$\begin{aligned} \underline{P}^1(f) &= \min\{P^1(f) : \forall P^1 \in ext(K(D^1))\} \\ &= \min\{ \left(\frac{2}{3} \times f(G) + 0 \times f(B) + \frac{1}{3} \times f(N)\right), \\ &\qquad \left(\frac{1}{2} \times f(G) + \frac{1}{6} \times f(B) + \frac{1}{3} \times f(N)\right), \\ &\qquad \left(\frac{1}{2} \times f(G) + 0 \times f(B) + \frac{1}{2} \times f(N)\right)\} \\ &= \min\{1.66, 1.33, 1.5\} = 1.33 \ . \end{aligned}$$

Similarly, the coherent upper prevision of gamble f is computed as follows:

$$\begin{aligned} \overline{P}^1(f) &= \max\{P^1(f) : \forall P^1 \in ext(K(D^1))\} \\ &= \max\{1.66, 1.33, 1.5\} = 1.66 \ . \end{aligned}$$

Replacing lower and upper previsions in Table 6 with their relevant values returns the enhanced GRR matrix in Table 7. As seen in Figure 3, the correlated equilibria polytope of this matrix is a heptahedron with six vertices, listed in Table 8. Vertices two and six are pure Nash equilibria that sit at the intersection between the polytope and the simplex representing all probability distributions on outcomes, *i.e.* the tetrahedron. The remaining vertices are correlated equilibria. It should be noted that the mixed Nash equilibrium of the original version of the game does not satisfy the correlated equilibria constraints of the modified

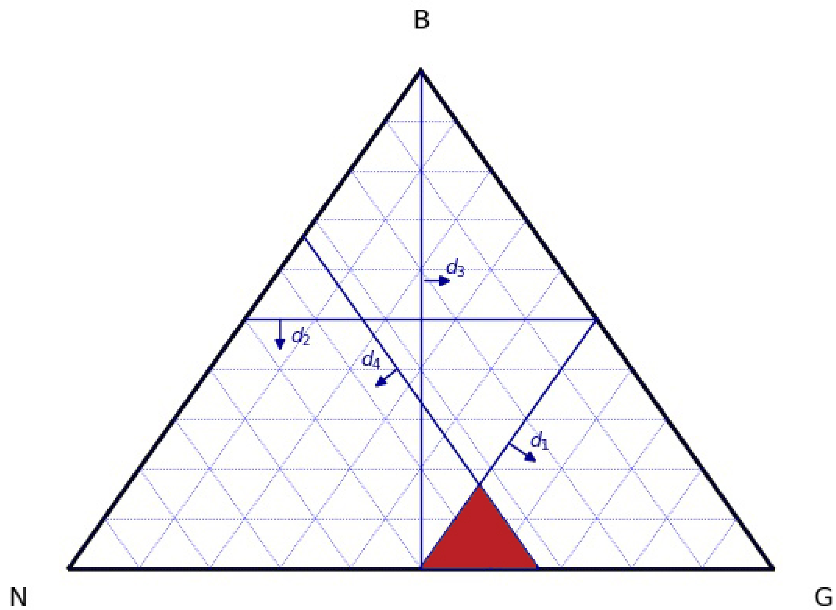

Figure 2: Example's resulting credal set.

Table 7: Example's resulting enhanced GRR matrix.

	TL	TR	BL	BR
$r^1_T - r^1_B$	1.33	-1	0	0
$r^1_B - r^1_T$	0	0	-1.66	1
$r^2_L - r^2_R$	1	0	-2	0
$r^2_R - r^2_L$	0	-1	0	2

one. Hence, on the inefficient frontier, the polytope does not intersect with the saddle that represents all joint probability distributions that are independent between players. That's because the supremum buying price and infimum selling price of f are different. This price mismatch shows that the GRR matrix reveals information that is not obvious by just looking at the payoff matrix. Especially under ignorance, when players might have two different buy and sell values for a specific payoff.

Optimal solutions for this game sit on the edge connecting TL and BR. Choosing one of them depends on the player's level of pessimism. For instance, an extremely pessimistic player would consider TL's payoff as the gamble's lower prevision, *i.e.* $r^1(TL) = \underline{P}^1(f) = 1.33$. This results in an optimal equilibrium $\rho_{TL} = 0.7518$ and $\rho_{BR} = 0.2482$, with an expected game payoff of 1.248 for both players. However, an extremely optimistic player would consider $r^1(TL) = \overline{P}^1(f) = 1.66$. This results in an optimal equilibrium $\rho_{TL} = 0.601$

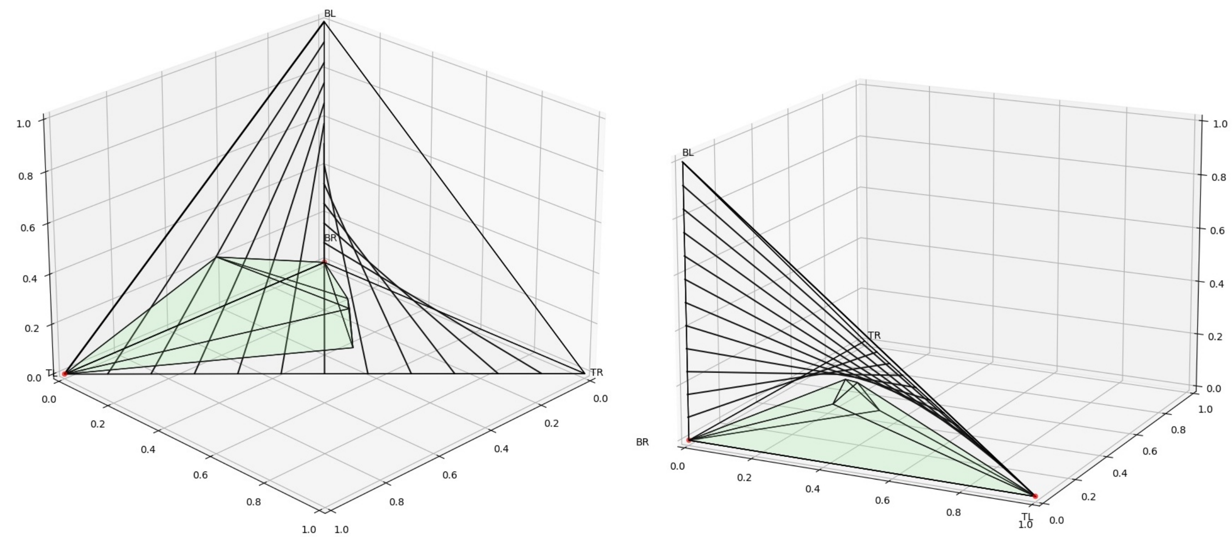

Figure 3: Example's first assessment polytope.

and $\rho_{BR} = 0.399$, with an expected payoff of 1.39 for both players.

Table 8: Example's first assessment vertices.

	TL	TR	BL	BR
Vertex 1	0.429	0	0.215	0.356
Vertex 2	1	0	0	0
Vertex 3	0.294	0.392	0.118	0.196
Vertex 4	0.334	0.444	0	0.222
Vertex 5	0.273	0.363	0.137	0.227
Vertex 6	0	0	0	1

4.2 Dynamic Updating

As discussed in Section 3.3, dynamic updating improves an uncertain outcome's lower and upper previsions in repeated games. Table 9 illustrates three different scenarios of applying dynamic updating to a sequence of plays. It shows that whenever outcome TL has a new payoff, the underlying gamble's lower and upper previsions are updated. Consequently, this triggers an update to the optimal correlated strategy.

Scenario one considers that the first time a play ends with an outcome TL, player one enjoys it more than all other outcomes. Hence, the payoff of gamble f is 2. Since this observation falls above the elicited $\overline{P}^1(f)$, the updated lower and upper previsions increase in value. Furthermore, $\underline{P}^1(f|u_1) = 1.665$ indicates that an extremely optimistic or pessimistic player expects TL's payoff to be the highest amongst all other payoffs still. If player one keeps getting a payoff of 2 for TL, the lower and upper previsions will eventually converge towards 2.

Table 9: Dynamic updating applied to three different scenarios.

n	Scenario 1			Scenario 2			Scenario 3		
	$u_n = f(x)$	$\underline{P}^1(f)$	$\overline{P}^1(f)$	$u_n = f(x)$	$\underline{P}^1(f)$	$\overline{P}^1(f)$	$u_n = f(x)$	$\underline{P}^1(f)$	$\overline{P}^1(f)$
0	–	1.33	1.66	–	1.33	1.66	–	1.33	1.66
1	2	1.665	1.83	0	0.665	0.83	0	0.665	0.83
2	2	1.776	1.886	0	–	–	2	1.11	1.22
3	2	1.832	1.915	0	–	–	2	1.332	1.415

Scenario two shows the opposite case. After experiencing outcome TL, player one considers it the worst outcome in the game. Hence, the payoff is 0. Since this observation falls below the elicited $\underline{P}^1(f)$, both updated previsions decrease in value. An upper prevision $\overline{P}^1(f|u_1) = 0.83$ is strictly smaller than the payoff player one gets from outcome BR. Therefore, an extremely optimistic or pessimistic player would stop choosing the alternative T and settles for a correlated strategy of $\rho_{TL} = 0$ and $\rho_{BR} = 1$, *i.e.* the pure Nash equilibrium. It should be noted that an extremely optimistic player would still consider the alternative T if the upper prevision is higher than BR's payoff. Scenario three shows the case where player one has different experiences related to outcome TL. The first time it is observed, its payoff is 0. This payoff decreases the upper prevision to $\overline{P}^1(f|u_1) = 0.83$, which is enough for a player to discard the alternative T. However, in practice, a player can still explore the outcome TL as a trial attempt. In that case, if the second observation has a payoff of 2, $\underline{P}^1(f|u_1)$ and $\overline{P}^1(f|u_2)$ increase in value and both become strictly higher than BR's payoff.

4.3 Extreme Ignorance

To illustrate our model under a case of extreme ignorance, we consider a variant of the game 'Matching Pennies'. The classic version is described by Von Neumann and Morgenstern (1944) as a game where two players simultaneously and independently select 'Heads' or 'Tails' each and then uncover a penny. If their selections match, then player two must give a penny to player one. Otherwise, player one gives a penny to player two. However, here, we modify the game so that player two gives player one an arbitrary reward generated by a gamble g. This gamble consists of drawing a ball from an urn. Depending on its colour, the following rewards are generated: 0 for red, 1 for black, and 2 for green.

Table 10: The modified version of 'matching pennies' - payoff matrix.

	T	H
T	$g, -g$	$-1, 1$
H	$-1, 1$	$g, -g$

We assume that information is symmetric across players throughout the gameplay and that no information about the composition of the urn is available to them. Therefore, vacuous lower and upper previsions are used; $\underline{P}^1(g) = \underline{P}^2(g) = 0$ and $\overline{P}^1(g) = \overline{P}^2(g) = 2$. Under these circumstances, whether the game is played or not depends on the adopted choice rule. For instance, using the pessimist/optimist decision rules, if any of the players is a pessimist, they will use the lower previsions of underlying gambles to assess expected payoffs. For outcomes TT and HH, they would only expect to lose utility when playing this game. Hence, they don't have any incentive to play it. However, if players one and two are optimists, they would expect, respectively, a payoff of $\overline{P}^1(g) = 2$ and $\overline{P}^2(-g) = -\underline{P}^2(g) = 0$ for both of these outcomes. Therefore, when no reward is yet observed for gamble g, the game's expectation for both optimists players is 0.5 and its mixed Nash equilibrium is (1/2 H, 1/2 T)×(1/2 H, 1/2 T).

Consider that the urn contains one black, one green, and two red balls. Applying the suggested dynamic updating algorithm to 1000 simulations of 200 plays each, returns an average lower prevision of $\underline{P}^1(g|u_{200}) = \underline{P}^2(g|u_{200}) = 0.74$ and an upper prevision of $\overline{P}^1(g|u_{200}) = \overline{P}^2(g|u_{200}) = 0.76$. Figure 4 shows how gamble g's lower and upper previsions

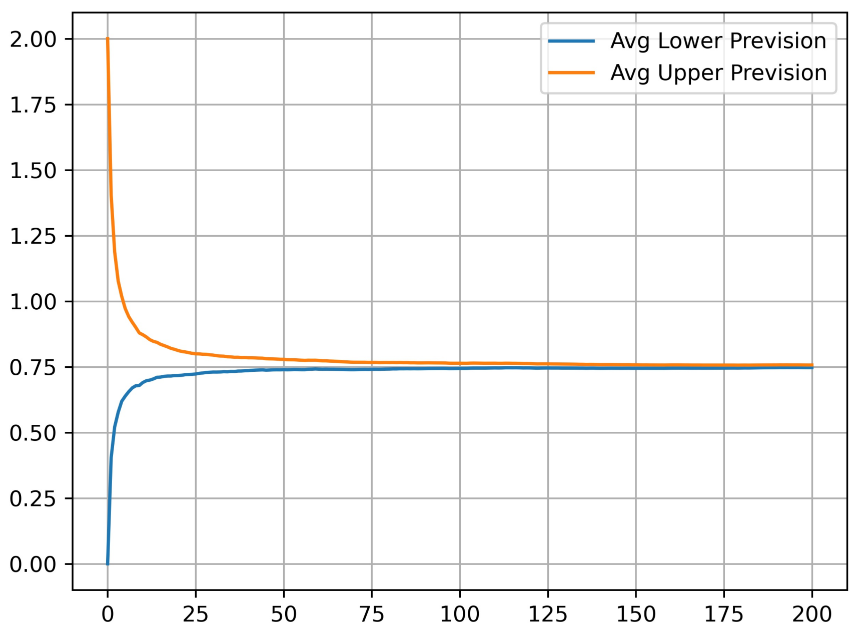

Figure 4: The modified version of 'matching pennies' - average lower and upper previsions generated using 1000 simulations of 200 plays each.

converge towards its linear prevision; that is, the actual expected payoff $P(g) = 0.75$, which is unknown to the players. The dynamic updating will influence each player's estimate of the game's expected payoff. As seen in Figure 5, on average, player one's expectation becomes negative after the ninth observation, giving them no incentive to keep playing the game. In contrast, player two's expectation is always positive. This indicates that player two has an advantage over player one, which is expected since the urn contains two red balls.

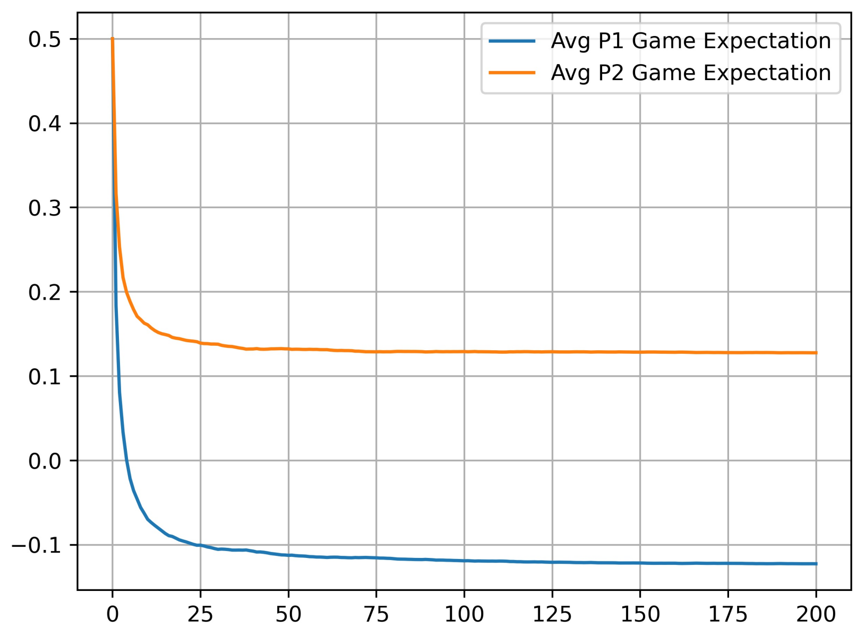

Figure 5: The modified version of 'matching pennies' - players' average game expectations generated using 1000 simulations of 200 plays each.

5 Discussion

The proposed solution could help expand the existing scope of application of game theory and allow it to include cases of ignorance over outcomes. Furthermore, we believe this solution could be extended to cooperative games. For instance, in oligopoly pricing (Athey and Bagwell, 2001) or repeated partnerships (Radner et al., 1986), players are assumed to know the utility of each outcome and the set of correlated distributions over these outcomes. However, this strong assumption could be relaxed using an enhanced GRR matrix. To illustrate, consider Radner's (1986) work on improving the decentralised decision-making process in an organisation. He studies repeated partnership games in which players cannot observe each other's strategies. In his example, two players contribute separate efforts to an enterprise. The combined effort of all players leads the enterprise to succeed or fail. Hence, $\mathcal{X} = \{\text{Success(S)} = 1, \text{Fail(F)}=0\ \}$. Players choose their effort simultaneously without being able to monitor each other's choices. The probability of success is considered $P(x = 1) = \min(e^1+e^2, 1)$, where e^1 and e^2 are the individual efforts of players one and two. Furthermore, the utility payoff of a player k is $f^k(x, e^k) = x - \lambda(e^k)^2$, where $\lambda > 0$, and λa^k represents the 'disutility of effort'. In this game, when a player contributes high effort, they know it is more likely to yield success. However, they don't know the exact probability of failure when all players choose to do the same. Therefore, players are uncertain about their expected payoffs. An enhanced GRR matrix can help assess this uncertainty and compute the set of correlated equilibria. Consider that a high effort(H) is when $e^k > 0.5$ and a low effort(L) is when $e^k \leq 0.5$. Let $\Lambda^1 = \Lambda^2 = \{\text{H, L}\}$ denote the alternatives available to each player. Hence, the set of possible outcomes is $\Phi = \{HH, HL, LH, LL\}$. Outcomes HH and LL represent respectively a sure success and a sure loss. However, both outcomes HL and LH, represent either a success or a loss. Assuming $\lambda = 1$, the payoff functions of these outcomes can be modelled as follows:

- $r^k(HH) = f_1^k(e^k) = 1 - (e^k)^2$, this function represents the payoff each player k gets when they and the other player choose to contribute high effort. Since for the outcome HH each player has to provide a minimum effort of 0.5, the lower and upper previsions of $r^k(HH)$ are $\underline{P}^k(r^k(HH)) = 0.75$ and $\overline{P}^k(r^k(HH)) = 0$ (assuming that the maximum effort is 1);

- $r^k(HL) = f_2^k(x, e^k) = x - (e^k)^2$ and $r^k(LH) = f_3^k(x, e^k) = x - (e^k)^2$, these functions represent the payoff each player k gets when they choose to contribute high effort while the other player contributes low effort, and vice versa. The lower and upper previsions of $r^k(HL)$ are $\underline{P}^k(r^k(HL)) = \underline{P}^k(x) - (\underline{P}^k(e^k))^2$ and $\overline{P}^k(r^k(HL)) = \overline{P}^k(x) - (\overline{P}^k(e^k))^2$. The same logic applies to payoff $r^k(LH)$. Hence, $\underline{P}^1(r^1(HL)) = \underline{P}^2(r^2(LH)) = 0.25$, $\overline{P}^1(r^1(HL)) = \overline{P}^2(r^2(LH)) = 0$, $\underline{P}^1(r^1(LH)) = \underline{P}^2(r^2(HL)) = 0.5$, and $\overline{P}^1(r^1(LH)) = \overline{P}^2(r^2(HL)) = 0.75$;

- $r^k(LL) = f_4^k(e^k) = -(e^k)^2$, this function represents the payoff each player k gets when they and the other player choose to contribute low effort. The lower and upper previsions of $r^k(LL)$ are $\underline{P}^k(r^k(LL)) = 0$ and $\overline{P}^k(r^k(LL))$= -0.25.

The resulting vacuous previsions should be refined using the provided elicitation model and then used in the enhanced GRR matrix. Given the payoff uncertainties, this matrix will return a convex set of correlated equilibria. Furthermore, in the repeated version of this game, dynamic updating can be used to adjust payoff expectations to the behaviour of each player. Dynamic updating will lead to an improved set of correlated equilibria by providing a more accurate utility assessment for each outcome.

6 Conclusions

The treatment of uncertain preferences over a game's outcome has had limited discussions in the Bayesian literature, even though such cases are empirically evident. Here, we proposed the enhanced GRR matrix complemented with an elicitation model and NPUI-based dynamic updating as a normative solution to those situations.

First, we developed an enhanced version of Nau's revealed-rules matrix that allows cases of uncertainty over a game's outcomes. This enhanced GRR matrix revealed information that is not observable within the game's payoff matrix. In particular, it allowed a player to disclose their beliefs about the lower and upper previsions of an uncertain outcome's payoff.

Second, we introduced an elicitation model to allow a player to coherently assess the lower and upper previsions of an uncertain outcome's payoff. This elicitation model can use pre-existing beliefs as a form of information. Hence, it can support cases of extreme ignorance, *i.e.* when a player has no information regarding an outcome.

Third, we computed the game's correlated equilibria using the enhanced GRR matrix enriched with players' relevant previsions.

Finally, we examined the case of repeated games and introduced a dynamic updating model. This model allows a player to adjust their previsions based on new observations. Specifically when experiencing uncertain outcomes.

In conclusion, we believe developing the proposed model further is possible. In particular, enhance it to address the more general case of risk-averse players. For instance, under the standard case, which doesn't involve any uncertainty, Nau (2011) shows that the game's rational solution is a convex set of equilibria whose parameters are risk-neutral probabilities. Therefore, under uncertainty, lower and upper previsions of each transaction, $r_i^k(\boldsymbol{\phi})-r_j^k(\boldsymbol{\phi})$, in $\boldsymbol{M}$ should be assessed with respect to a convex set of risk-neutral probabilities. Furthermore, it is possible to explore the application of the proposed model to different fields of study. For example, within Artificial Intelligence, if a payoff matrix contains unobserved outcomes, it could be replaced with an enhanced GRR matrix that includes players' elicited lower and upper previsions towards these outcomes. Or within Economics, in rivalry or alliance situations where ignorance can prevail. For instance, when companies invest in a new market and unforeseen events arise, this could cause a conflict of interest. In this case, stakeholders could update their preferences using the dynamic updating mechanism.

References

Astanin, S. V. and Zhukovskaja, N. K. (2015). Using game theory and fuzzy logic to determine the dominant motivation cognitive agent. *American Association for Science and Technology*, 2:207–214.

Athey, S. and Bagwell, K. (2001). Optimal collusion with private information. *RAND Journal of Economics*, 32:428–465.

Augustin, T. and Coolen, F. (2004). Nonparametric predictive inference and interval probability. *Journal of Statistical Planning and Inference*, 124(2):251 – 272.

Aumann, R. J. (1974). Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1(1):67 – 96.

Aumann, R. J. (1976). Agreeing to disagree. *The Annals of Statistics*, 4(6):1236–1239.

Aumann, R. J. (1987). Correlated equilibrium as an expression of bayesian rationality. *Econometrica*, 55(1):1–18.

Chakeri, A., Dariani, A. N., and Lucas, C. (2008). How can fuzzy logic determine game equilibriums better? In *2008 4th International IEEE Conference Intelligent Systems*, pages 100–105. IEEE.

Chiandotto, B. (2014). *Bayesian and Non-Bayesian Approaches to Statistical Inference: A Personal View*, pages 3–13. Springer International Publishing, Cham.

Chiara, M., Doets, K., Mundici, D., and van Benthem, J. (2013). *Structures and Norms in Science: Volume Two of the Tenth International Congress of Logic, Methodology and Philosophy of Science, Florence, August 1995*. Synthese Library. Springer Netherlands.

de Finetti, B. (1937). La prévision: Ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 17:1–68.

de Finetti, B. (1974, 1975). *Theory of Probability: A Critical Introductory Treatment*, volume 1 and 2 of *Probability and Statistics Series*. Wiley.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 75(4):643–669.

Gajdos, T., Tallon, J. M., and Vergnaud, J. C. (2004). Decision making with imprecise probabilistic information. *Journal of Mathematical Economics*, 40(6):647–681.

Gale, D. (1960). *The Theory of Linear Economic Models*. McGraw-Hill, New York.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Harsanyi, J. (1967). Games with incomplete information played by "bayesian" players, i-iii part i. the basic model. *Management Science*, 14(3):159–182.

Hill, B. (1968). Posterior distribution of percentiles: Bayes' theorem for sampling from a population. *Journal of the American Statistical Association*, 63(322):677–691.

Hill, B. (1988). De Finetti's theorem, induction, and $A_{(n)}$ or Bayesian nonparametric predictive inference (with discussion). *Bayesian statistics*, 3:211–241.

Hill, B. (1993). Parametric models for $A_{(n)}$: Splitting processes and mixtures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(2):423–433.

Houlding, B. and Coolen, F. (2012). Nonparametric predictive utility inference. *European Journal of Operational Research*, 221:222–230.

Jansen, C., Blocher, H., Augustin, T., and Schollmeyer, G. (2022). Information efficient learning of complexly structured preferences: Elicitation procedures and their application to decision making under uncertainty. *International Journal of Approximate Reasoning*, 144:69–91.

Luce, R. and Raïffa, H. (1957). *Games and Decisions: Introduction and Critical Survey*. Publications of the Bureau of applied social research of the Columbia University. John Wiley & Sons, Incorporated.

Nash, J. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49.

Nau, R. (1992). Joint coherence in games of incomplete information. *Management Science*, 38(3):374–387.

Nau, R. (2011). Imprecise probabilities in non-cooperative games. *ISIPTA 2011 - Proceedings of the 7th International Symposium on Imprecise Probability: Theories and Applications*.

Nau, R., Canovas, S., and Hansen, P. (2004). On the geometry of nash equilibria and correlated equilibria. *International Journal of Game Theory*, 32:443–453.

Nau, R. and McCardle, K. (1990). Coherent behavior in noncooperative games. *Journal of Economic Theory*, 50(2):424 – 444.

Pareto, V. and Bonnet, A. (1963). *Manuel d'Économie Politique*. Number v. 2 in Bibliothèque internationale d'économie politique. Librairie générale de droit et de jurisprudence.

Radner, R. (1986). Repeated Partnership Games with Imperfect Monitoring and No Discounting. *The Review of Economic Studies*, 53(1):43–57.

Radner, R., Myerson, R., and Maskin, E. (1986). An example of a repeated partnership game with discounting and with uniformly inefficient equilibria. *Review of Economic Studies*, 53:59–70.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, 57:571–87.

Troffaes, M. C. (2007). Decision making under uncertainty using imprecise probabilities. *International Journal of Approximate Reasoning*, 45(1):17–29.

Von Neumann, J. and Morgenstern, O. (1944). *Theory Of Games And Economic Behavior*. Princeton University Press.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall.

Williams, P. M. (1974). Indeterminate probabilities. In Przelecki, M., Szaniawski, K., Wójcicki, R., and Malinowski, G., editors, *Formal Methods in the Methodology of Empirical Sciences: Proceedings of the Conference for Formal Methods in the Methodology of Empirical Sciences, Warsaw, June 17–21, 1974*, pages 229–246, Dordrecht, Netherlands.

Williams, P. M. (1975). Coherence, strict coherence and zero probabilities. In *DLMPS '75: Proceedings of the Fifth International Congress of Logic, Methodology and Philosophy of Science, volume VI*, pages 29–33.