

Incorporating Belief-Dependent Motivation in Games*

Pierpaolo Battigalli[†]

Roberto Corrao[‡]

Martin Dufwenberg[§]

3 February 2019

Abstract

Psychological game theory (PGT), introduced by Geanakoplos, Pearce & Stacchetti (1989) and significantly generalized by Battigalli & Dufwenberg (2009), extends the standard game theoretic framework by letting players' utility at endnodes depend on their interactive beliefs. While it is understood that a host of applications that model and/or test the role of emotional and other psychological forces find their home in PGT, the framework is abstract and comprises complex mathematical objects, such as players' infinite hierarchies of beliefs. Thus, PGT provides little guidance on how to model specific belief-dependent motivations and use them in game theoretic analysis. This paper takes steps to fill this gap. Some aspects are simplified – e.g., which beliefs matter – but others are refined and brought closer to applications by providing more structure. We start with belief-dependent motivations and show how to embed them in game forms to obtain psychological games. We emphasize the role of time and of the perception of players' intentions. We take advantage of progress made on the foundations of game theory to expand and improve on PGT solution concepts.

JEL classification: C72; C73; D81; D82; D92

Keywords: Psychological game theory; Belief-dependent motivation; Intentions; Time; Rationalizability; Self-confirming equilibrium; Bayesian sequential equilibrium

1 Introduction

Following Elster (1998), economists have become increasingly aware that belief-dependent motivation is important to decision making, and that this can have important economic consequences. Caplin & Leahy (2001, 2004), for instance, propose a model where anxiety depends on the degree of uncertainty of an agent's beliefs about his future health or wealth, influencing his utility & behavior as well as how (concerned) others treat him.¹ These phenomena cannot be examined using

*We have benefited from helpful comments from or discussions with Giuseppe Attanasi, Simone Cerreia-Vioglio, Carlo Cusumano, Francesco Fabbri, Rachel Mannahan, Elena Manzoni, Paola Moscarrello, Amrish Patel, Alec Smith, and three referees. Financial support of ERC (grant 324219) and of the Marco Fanno scholarship are gratefully acknowledged.

[†]Bocconi University and IGIER, Italy: pierpaolo.battigalli@unibocconi.it

[‡]MIT, USA: rcorrao@mit.edu

[§]U of Arizona, USA; U of Gothenburg, Sweden; CESifo, Germany: martind@eller.arizona.edu

¹A handful of other examples concern status & conformity (Bernheim 1994), reciprocity & MOUs (Jang et al. 2018), guilt & tax evasion (Dufwenberg & Nordblom 2018), disappointment & savings (Koszegi & Rabin 2009), anger & bargaining (Battigalli et al. 2015, Aina et al. 2018, Dufwenberg et al. 2018a,b), and deception (Battigalli et al. 2013, Dufwenberg & Dufwenberg 2018, Gneezy et al. 2018).

traditional game theory where utility depends only on material outcomes, hence on choices. One rather needs the mathematical framework introduced and labeled **psychological game theory** (PGT) by Geanakoplos et al. (1989) (GP&S) and further developed by Battigalli & Dufwenberg (2009) (B&D). The key objects, called **psychological games** (p-games), can describe a variety of belief-dependent motivations including emotions, reciprocity, self-esteem, and caring about others' opinions.

As developed so far, PGT is quite abstract, leaving implicit or unexplored important links between psychological assumptions and economic outcomes, on the one hand, and mathematical modeling on the other. It may not be easy for scholars doing applied work to translate psychological and economic notions of interest into the formal framework, and it is not obvious how to derive “solutions.” Furthermore, little has been said concerning the relationship between motivation and time. For instance, in the case of frustration & anger, the phenomenon of cooling-off, whereby anger subsides over time, is well known among psychologists, hence it should be explicitly modeled.

We provide a novel methodology to embed belief-dependent motivations in strategic analysis. Our approach has two parts, generating p-games à la B&D and providing predictions:

The first part (=sections 2-6) We add psychological preferences to the rules of the game. The latter are represented by game forms, providing an explicit role for time, and distinguishing periods from stages. The former have enough duration to be relevant for discounting, decay, cooling-off, expectation-based reference points, etc.; the latter allow for modelling the sequential moves taking place close in time, during each period, which in turn give rise to material outcomes. We explicitly model incomplete and asymmetric information about personal traits (including psychological ones),² and some features of the environment.

The next key step addresses agents' intertemporal psychological preferences. Absent belief-dependence, utility can be defined over temporal sequences (streams) of material outcomes. To allow for belief-dependent motivations we let utility depend also on temporal sequences of beliefs (own and others') about *outcomes and personal traits*. Many forms of preferences (e.g. anxiety, disappointment, simple guilt) depend only on those sequences, and can then be modeled independently of the game form in which they are embedded, much like standard preferences over lotteries and intertemporal preferences. We describe the method to thus obtain a p-game in detail. To emphasize that the involved preferences can be discussed without reference to any specific game form, we label them **game-independent**, and present them before our formal representation of the rules of the game.

Other forms of belief-dependent motivation—e.g. reciprocity, regret, anger, perceived cheating aversion—are game-dependent. They refer to players' behavior, intentions, or inferences in particular game forms, and require more machinery. We model them considering players' beliefs about (traits and) *actions*, which we call **first-order beliefs**. Their description is key to our analysis. They entail beliefs also of the sort described in the previous paragraph, but the added focus on actions is central for describing **game-dependent** motivation. We let psychological utility depend on temporal sequences of first-order beliefs. In principle, beliefs about beliefs—higher-

²Personal traits are exogenous personal features, such as general ability, or *personality* traits such as the “big five” (see Rothmann & Coetzer 2003 and references therein).

order beliefs—may matter to utility too (e.g., Battigalli & Dufwenberg 2007; guilt-from-blame), but we focus on utilities that depend on first-order beliefs (own & others’) because this simplifies much and it is still enough to cover most applications.

In so far as a player i ’s utility depends on the first-order beliefs of a co-player j , to compute the subjective value of his actions and rank them, i still has to consult his **second-order beliefs**, i.e., his beliefs about j ’s first-order beliefs. The expectation of belief-dependent utility according to second-order beliefs can be interpreted as a psychological utility “experienced” by the players; we refer to it (nodding to the spirit of Kahneman 1994 and Kahneman et al. 1997) as **experience utility**.

The maximization of expected experience utility covers many important applications and can also account for interesting forms of dynamic inconsistency (e.g., Section 6 of B&D; Koszegi & Rabin 2006, 2009); but some relevant action tendencies—such as the angry reaction of a frustrated agent—are not easily accommodated. Does this mean that we must abandon the maximization paradigm? We argue that this is not necessary. Some action tendencies such as anger and reciprocity can be accommodated by defining the expected **decision utility** of an action at a node as the sum of the expected experience utility associated with that action and the expected value of a “distortion.” For example, following Battigalli et al.’s (2015) (BD&S) work on frustration and anger in games, if the desire to maximize some kind of hedonic utility clashes with desire to punish blameworthy co-players (where who is blameworthy may depend on beliefs), the distortion could be proportional to how much money is taken away, either destroyed or given to others. The choice at the given node is then explained as the maximization of expected decision utility, which may differ from maximization of expected experience utility.

The second part (=section 7) We turn to predictions. Our goal is to develop better founded solutions. Most importantly, we believe it is high time to complement the usual approach of relying on standard equilibrium concepts (which assume—without serious justification—that players’ beliefs are correct) with other approaches.

We first define and examine subjective **rationality**, and argue that the notion should be understood via properties of a player’s beliefs, which can be decomposed into a **plan**—i.e., beliefs about own behavior—and beliefs about co-players.

Second, we put forward an operational version of **(extensive-form) rationalizability**, extending work by Pearce (1984) and B&D, and using some results from the epistemic analysis of Battigalli et al. (2019) (BC&S).³ Our new formulation is more flexible as we allow for some given and transparent restrictions on players’ beliefs. Rationalizability characterizes the behavioral and low-order belief implications of rationality and common strong belief in rationality, without any presumption of equilibrium.

Third, we turn to a setting with recurrent play, where players may learn over time and adjust behaviors. No presumption of equilibrium is made, but the implied forms of behavior satisfy an equilibrium property. The relevant notion, extended to and stated for p-games for the first

³Also Jagau & Perea (2017) and Bjorndahl et al. (2013) provide an epistemic analysis of rationalizability in p-games (Bjorndahl et al. focus on dependence of utility on coarse features of beliefs). They consider one-stage games where psychological utility depends only on initial beliefs while BC&S give the epistemic foundation of strong (Δ)-rationalizability as defined in this paper for dynamic psychological games.

time, is **self-confirming equilibrium** (SCE). It is the appropriate concept whenever one aims to analyze stable patterns of behavior in empirical data.⁴

Only thereafter do we turn to traditional equilibrium analysis. We argue that the extensions to PGT of traditional concepts, à la GP&S and B&D, should be re-examined. Developing ideas of Attanasi et al. (2016) and BD&S, we propose a new extension of Kreps & Wilson’s (1982) (K&W) **sequential equilibrium** (SE) concept called **Bayesian SE** (BSE). We point out that there is no need to provide *ad hoc* extensions of traditional solution concepts for PGT: if we accept the (very questionable) assumption that players’ conjectures about opponents’ decision rules are correct, then it is enough to carry out the logic of Harsanyi’s (1967-68) Bayesian equilibrium concept, which “mechanically” yields the endogenous hierarchies of beliefs that enter (expected) psychological utility.

Comparison with B&D Our approach is technically mainly consistent with B&D, but involves three innovations. First, B&D take as primitive utility functions that depend on the terminal node and players’ hierarchies of beliefs about behavior. Instead, we start from belief-dependent motivations and embed them in a game form to obtain a dynamic p-game, taking explicit account of time. Second, B&D rely on the assumption that the objective description of players’ contingent behavior coincides with their plans and that this is transparent,⁵ which prevents the possibility to define events like “player i planned to take action a but instead took action b ”. Unlike B&D, we distinguish plans from actual behavior, allowing us to model players who care about co-players’ intentions. Third, B&D allow utility to depend on beliefs of arbitrarily high order, whereas we simplify and consider only the first and second order. This allows for an algorithmic characterization of rationalizability, absent in B&D.

Rest of paper Section 2 provides a heuristic introduction to our framework. Section 3 gives the basic ingredients for the analysis of belief-dependent motivation developed in Section 4. Section 5 introduces game-forms and systems of conditional beliefs. Section 6 defines our notion of dynamic p-game. Section 7 analyzes solution concepts. Finally, Section 8 offers concluding remarks.

2 Heuristic examples

This section provides a gentle introduction to our conceptual framework and formalism, illustrating via examples the relation between psychological preferences and rules of the game. Consider the interaction between two agents, Ann & Bob (or A & B). A feasible allocation of monetary payoffs is a pair, or profile, $\mathbf{y} = (y_A, y_B) \in \mathbf{Y} \subseteq \mathbb{R}^{\{A,B\}}$. For now, we focus on a one-period interaction, ruling out intertemporal phenomena like discounting, decay, or cooling-off. Yet, within the single period, players may move sequentially.

⁴As regards SCE in standard games, see Fudenberg & Levine (1993) and the survey by Battigalli et al. (1992).

⁵B&D write (p. 11): “our specification of the conditioning events relies on interpreting [j ’s strategy] s_j as an objective description of how j would behave at each decision node. However, we will also interpret s_j as a *plan* in the mind of player j . The implicit assumption [...] is that each player has correct beliefs, given by his plan of action, about how he would choose at different histories”.

Let us briefly describe the notation we use for probability measures and expectations. Given an arbitrary measurable space (X, \mathcal{F}) , we denote the set of probability measures defined on \mathcal{F} with $\Delta(X)$, where the sigma-algebra is always taken as understood. Given a probability measure $\mu \in \Delta(X)$ and an integrable real function $f : X \rightarrow \mathbb{R}$, we denote by $E[f, \mu]$ the expectation of f with respect to μ . When $x \in X \subseteq \mathbb{R}$ and f coincides with the identity function, we write $E[\tilde{x}, \mu]$ for the expectation over X with respect to μ .

2.1 Game-independent psychological utilities

Each player $i \in \{A, B\}$ starts with an initial belief $\mu_{i,0} \in \Delta(\mathbf{Y})$ about material outcomes and ends the interaction with a terminal realized belief $\mu_{i,1} \in \Delta(\mathbf{Y})$ that depends on what i observed.⁶ We allow i 's utility to depend, in principle, on the initial and terminal beliefs of both Ann and Bob, i.e., on the following temporal sequence of belief profiles about \mathbf{Y} :

$$\boldsymbol{\mu} = (\boldsymbol{\mu}_0, \boldsymbol{\mu}_1) = \begin{pmatrix} \mu_{A,0} & \mu_{A,1} \\ \mu_{B,0} & \mu_{B,1} \end{pmatrix} \in [\Delta(\mathbf{Y})]^{\{A,B\} \times \{0,1\}}.$$

Thus, the abstract form of i 's utility is

$$v_i : \mathbf{Y} \times [\Delta(\mathbf{Y})]^{\{A,B\} \times \{0,1\}} \rightarrow \mathbb{R}.$$

From now on, assume that Ann holds standard preferences over material outcomes. Bob, however, has belief-dependent motivations, which, for now, we assume to be commonly known. Later on we remove this assumption. Note, following B&D, that we allow Bob's utility to depend on something he cannot observe: Ann's beliefs. This is in the spirit of the state-dependent utility functions often used in standard economics. Of course, to determine his subjectively optimal behavior, Bob has to consult his beliefs over the beliefs of Ann.

The nature of Bob's utility will vary depending on the sentiment that affects him. First, assume it depends negatively on Ann's disappointment (Battigalli & Dufwenberg 2007 call this "simple guilt"),⁷ defined by the difference, if positive, between the payoff Ann initially expected to receive and her actual payoff at the end of the period. Formally, players' utilities are given by

$$\begin{aligned} v_A(\mathbf{y}, \boldsymbol{\mu}) &= y_A, \\ v_B(\mathbf{y}, \boldsymbol{\mu}) &= y_B - \theta_B D_A[y_A, \mu_{A,0}] := y_B - \theta_B \max\{E[\tilde{y}_A, \mu_{A,0}] - y_A, 0\}, \end{aligned} \tag{1}$$

where θ_B is a personal trait of Bob measuring how much money he would be willing to give up to reduce Ann's disappointment by \$1. With this, we can embed the psychological preferences just described in any situation of one-period strategic interaction. Such situations, determined by the rules of the game, are called **game forms**. Recall that even one-period game forms may have a multistage structure. Technically, a game form comprises a game tree, specifying a set

⁶If i could either directly observe, or indirectly infer the realized outcome \mathbf{y} , then his terminal belief about the outcome would be trivial; this fits many situations we study. But we are also going to consider situations where \mathbf{y} is not perfectly observable, or i also cares about unknown personal traits of his co-players, which are only imperfectly revealed by their actions. See the end of this subsection for a simple illustrative example.

⁷See also Huang & Wu (1994), Dufwenberg & Gneezy (2000), Dufwenberg (2002), Charness & Dufwenberg (2006).

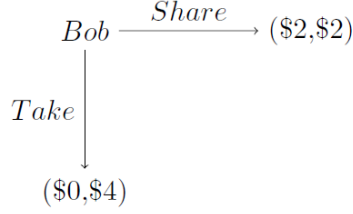


Figure 1: Dictator Minigame with material payoffs

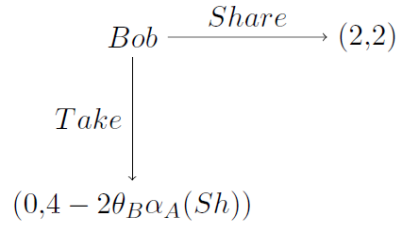


Figure 2: Dictator Minigame with psy-utility (guilt)

Z of complete paths of the game, and an **outcome function** $\pi : Z \rightarrow \mathbf{Y}$ that associates each complete path z (equivalently, terminal node) with a corresponding profile of material outcomes: $(y_A, y_B) = \pi(z) = (\pi_A(z), \pi_B(z))$.

We assume that the game form is commonly known. Players' utility functions are notably absent from game forms. Thus (somewhat confusingly and unfortunately) game forms correspond to the notion of “game” in the natural language: a set of rules, not the personal traits of the agents who happen to be players of the game. Of course, such personal traits are crucial, but—for conceptual clarity—we analyze them separately.

Fig. 1 depicts a Dictator Minigame.⁸ In its unique stage, the only active player is Bob, who can either *Take* (or *Tk*) \$4 or *Share* (*Sh*) that amount evenly. Even if Ann is not active, at the beginning (root) of the game she has a first-order belief $\alpha_A \in \Delta(\{Tk, Sh\})$ over Bob's choices. Let $\alpha_A(Sh) \in [0, 1]$ denote the subjective probability assigned by Ann to Bob choosing *Share*. Ann's initial beliefs about outcomes are easily derived:

$$\mu_{A,0}(\$2, \$2) = \alpha_A(Sh) \quad \text{and} \quad \mu_{A,0}(\$0, \$4) = 1 - \alpha_A(Sh) = 1 - \mu_{A,0}(\$2, \$2). \quad (2)$$

Embedding the utilities of Eq. (1) and using (2) we obtain the p-game of Fig. 2.

Bob consults his second-order belief in order to attach expected value to his actions: let $\bar{\beta}_B \in [0, 1]$ denote Bob's subjective expectation of $\alpha_A(Sh)$ (which he does not know). We obtain the alternative representation of Fig. 3, where each endnode is associated with Bob's

⁸Note, we give names like “Dictator Minigame” to game forms (usually with monetary outcomes), not to games with (standard, or psychological) utility functions.

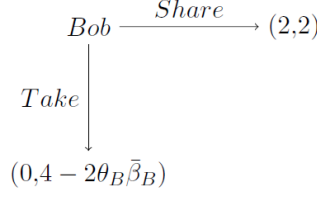


Figure 3: Dictator Minigame with exp-utility (guilt)

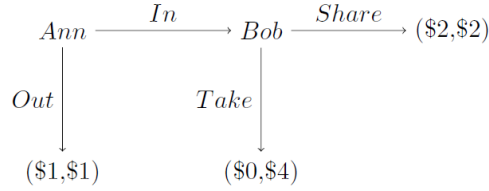


Figure 4: Trust Minigame with material payoffs

“experienced” utility, which is his subjectively expected psychological utility (exp-utility) given his action. *Share* is a best-reply for Bob if and only if $\theta_B \bar{\beta}_B \geq 1$.

The utilities of Eq. (1) can be embedded in any other game form. Consider, e.g., the Trust Minigame in Fig. 4, a two-stage game. Ann’s beliefs about outcomes now depend on her plan to “Trust” (or not) and her expectation of Bob’s reply. We represent both these expectations as (first-order) beliefs of Ann about how the game will be played, namely, a system of conditional (and unconditional) beliefs about actions

$$(\alpha_A(\cdot|\emptyset), \alpha_A(\cdot|In)) \in \Delta(\{In, Out\}) \times \Delta(\{Tk, Sh\}),$$

where Ann’s belief at the root (denoted \emptyset) about her own action represents her **plan**. Note, Ann’s plan is logically distinct from her behavior. We say that she is **consistent** if she carries out her plan. Analogous considerations apply to Bob.

We derive Ann’s initial beliefs about outcomes from her beliefs about behavior as follows:

$$\begin{aligned}
\mu_{A,0}(\$1, \$1) &= 1 - \alpha_A(In|\emptyset), \\
\mu_{A,0}(\$2, \$2) &= \alpha_A(Sh|In) \alpha_A(In|\emptyset), \\
\mu_{A,0}(\$0, \$4) &= \alpha_A(Tk|In) \alpha_A(In|\emptyset) = (1 - \alpha_A(Sh|In)) \alpha_A(In|\emptyset).
\end{aligned}$$

We use obvious abbreviations like $\alpha_A^{In} := \alpha_A(In|\emptyset)$ and $\alpha_A^{Sh} := \alpha_A(Sh|In)$. When we embed the psychological utilities of Eq. (1) in the Trust Minigame we obtain the p-game in Fig. 5, which assumes that Ann is commonly known to be selfish (and risk neutral). Unlike the game in Fig. 2, Bob’s utility depends on Ann’s plan on top of her beliefs about Bob’s choice. In particular,

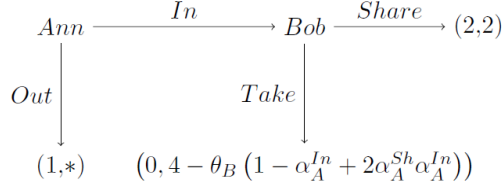


Figure 5: Trust Minigame with psy-utility (guilt)

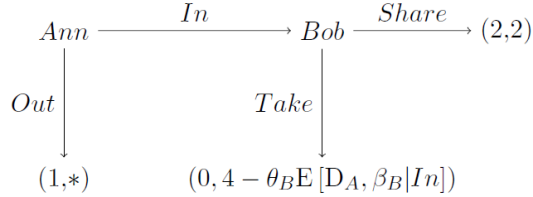


Figure 6: Trust Minigame with exp-utility (guilt)

given the equations relating μ_A with α_A , Ann’s disappointment at terminal history $z = (In, Tk)$ is

$$\begin{aligned}
D_A(\$0, \mu_{A,0}) &= \max \{E[\tilde{y}_A, \mu_{A,0}] - 0, 0\} = E[\tilde{y}_A, \mu_{A,0}] \\
&= 1 \times (1 - \alpha_A^{In}) + \left[2 \times \alpha_A^{Sh} + 0 \times (1 - \alpha_A^{Sh}) \right] \times \alpha_A^{In}.
\end{aligned}$$

For simplicity, here (and similarly later) we do not specify Bob’s psychological utility given *Out* as it is irrelevant for his decision.

Letting $E[D_A, \beta_B | In]$ denote Bob’s subjective expectations of Ann’s disappointment according to his *updated* second-order belief $\beta_B(\cdot | In)$, we obtain the alternative representation of Fig. 6, where each endnode is associated with Bob’s experienced utility. *Share* is a best-reply for all (updated) second-order beliefs $\beta_B(\cdot | In)$ such that $\theta_B E[D_A, \beta_B | In] \geq 2$. For example, if Bob is certain, upon observing “In,” that Ann truly planned to play “In” (an assumption that we do *not* take for granted, see Section 7), then $E[D_A, \beta_B | In] = 2\bar{\beta}_B^{Sh}$, where $\bar{\beta}_B^{Sh}$ denotes Bob’s subjective expectation of Ann’s first-order belief.⁹ With this, we obtain the same condition as in the Dictator Minigame: $\theta_B \bar{\beta}_B^{Sh} \geq 1$.

In general, we allow for incomplete information, i.e., that players’ utilities and outcome functions depend on a parameter $\theta \in \Theta$ which is *not* perfectly known to all. Players may have different exogenous information regarding θ . Each i forms beliefs about paths $z \in Z$ and about θ given private information and conditional on reaching each node (partial or complete path) of

⁹Indeed, given that Bob has a degenerate belief concerning Ann’s plan, from his point of view the random variables $\tilde{\alpha}_A^{In}$ and $\tilde{\alpha}_A^{Sh}$ are stochastically independent.

the game tree. Such beliefs include i 's plan how to play and form i 's **system of first-order (conditional) beliefs**, denoted α_i . In the previous examples, we informally referred to the first-order conditional beliefs of Ann in order to derive her plan and her initial beliefs over profiles of monetary payoffs. Formally, systems of first-order conditional beliefs are defined as maps from histories (nodes) to probability distributions on complete paths (terminal nodes) and θ . As we show in Section 5, for each i and z , one can derive from α_i a temporal sequence of **realized** beliefs about behavior, outcomes, and θ . With complete information, terminal beliefs are point beliefs that assign probability 1 to outcomes. This is not the case with incomplete information.

To illustrate, we embed different utilities in the game of Fig. 1. Drop the assumption of complete information and consider a finite set of parameters $\Theta_B := \Theta_B^{al} \times \Theta_B^{rep} \subseteq \mathbb{R}_+^2$ spanning Bob's personal traits: θ_B^{al} measures Bob's altruism while θ_B^{rep} measures a "reputation concern," i.e., a concern for the opinion of Ann about his altruism (cf. Ellingsen & Johannesson 2008). Since $(\theta_B^{al}, \theta_B^{rep})$ is unknown to Ann, her initial beliefs and terminal realized beliefs will be defined on the product space of personal traits of Bob and monetary payoffs, i.e., $\Theta_B \times \mathbf{Y}$. The sequence of belief profiles is a vector $\boldsymbol{\mu} = ((\mu_{A,0}, \mu_{A,1}), (\mu_{B,0}, \mu_{B,1}))$, where $\mu_{A,0}, \mu_{A,1} \in \Delta(\Theta_B \times \mathbf{Y})$, $\mu_{B,0}$ ($\$2, \2) is the probability with which Bob's plans to *Share*, and $\mu_{B,1}$ is trivial. The players' utility functions are as follows:

$$\begin{aligned} v_A(\mathbf{y}, \boldsymbol{\mu}) &= y_A, \\ v_B(\theta_B, \mathbf{y}, \boldsymbol{\mu}) &= y_B + \theta_B^{al} y_A + \theta_B^{rep} \mathbb{E} \left[\tilde{\theta}_B^{al}, \mu_{A,1} \right], \end{aligned}$$

where $\mathbb{E} \left[\tilde{\theta}_B^{al}, \mu_{A,1} \right]$ is Ann's terminal estimate of Bob's altruism. Thus, θ_B^{rep} measures how much money Bob would be willing to give up to increase Ann's estimate of his altruism by one unit. The main difference with respect to Bob's psychological utility of Eq. (1) is that here he cares about the terminal beliefs of Ann.

To embed the modified psychological utility of Bob in the Dictator Minigame we need to specify the entire system of conditional first-order beliefs of Ann. This is a function

$$\alpha_A : \{\emptyset, Sh, Tk\} \rightarrow \Delta(\{Sh, Tk\} \times \Theta_B)$$

that associates each history in $\{\emptyset, Sh, Tk\}$ with a belief about paths and Bob's personal traits. Of course, α_A cannot be exogenously given; it is determined by Ann's strategic reasoning. Furthermore, at least one of the two terminal beliefs $\alpha_A(\cdot|Sh)$ and $\alpha_A(\cdot|Tk)$ is pinned down by the initial belief $\alpha_A(\cdot|\emptyset)$ and the rules of conditional probabilities.¹⁰ For each $z \in \{Sh, Tk\}$, Ann's system of first-order beliefs α_A determines $(\mu_{A,0}, \mu_{A,1})$ as follows:

$$\begin{aligned} \mu_{A,0} \left(\left(\theta_B^{al}, \theta_B^{rep} \right), \mathbf{y} \right) &= \alpha_A \left(\boldsymbol{\pi}^{-1}(\mathbf{y}), \left(\theta_B^{al}, \theta_B^{rep} \right) | \emptyset \right), \\ \mu_{A,1} \left(\left(\theta_B^{al}, \theta_B^{rep} \right) \right) &= \alpha_A \left(\theta_B^{al}, \theta_B^{rep} | z \right) \end{aligned}$$

for every $((\theta_B^{al}, \theta_B^{rep}), \mathbf{y}) \in \Theta_B \times \mathbf{Y}$.¹¹ For example, suppose that $\Theta_B = \{0, \bar{\theta}_B^{al}\} \times \{0, \bar{\theta}_B^{rep}\}$ and

¹⁰We use obvious abbreviation for marginal probabilities. Both probabilities are determined by $\alpha_A(\cdot|\emptyset)$ if $0 < \alpha_A(\{Sh\} \times \Theta_B | \emptyset) < 1$.

¹¹In the Dictator Minigame, and in all the simple examples considered, the outcome function $\boldsymbol{\pi} : Z \rightarrow \mathbf{Y}$ is injective and therefore, for every $\mathbf{y} \in \boldsymbol{\pi}(Z)$, $\boldsymbol{\pi}^{-1}(\mathbf{y})$ is a well defined terminal path in Z .

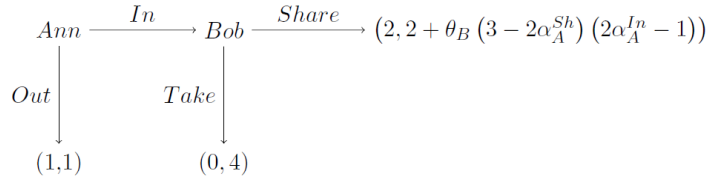


Figure 7: Trust Minigame with psy-utility (reciprocity)

that Ann (reasonably) believes that Bob is more likely to *Share* if θ_B^{al} or θ_B^{rep} are high. Then $\mu_{A,1} \left(\left(\bar{\theta}_B^{al}, \bar{\theta}_B^{rep} \right) \right)$ is larger after *Share* than after *Take*. It follows that Bob has a greater incentive to *Share*.

2.2 Game-dependent psychological utilities

The previous examples tempt us to try and describe *all* the possible belief-dependent motivations with psychological utilities that do not depend on the game situation. This would be similar to the standard approach in economic theory where attitudes toward risk and the timing of consumption are modeled as personal traits of agents, and then embedded in the games they play. However, some forms of belief-dependent motivations depend on the players' feasible choices. Since those are described by the game form, there is a class of psychological preferences which cannot be formally represented without reference to a given game form.

For example, in reciprocity theory, the kindness of an action is assessed with reference to the (expected) payoff distributions associated with other feasible actions, given by the game form (and players' beliefs). We illustrate this with the game of Fig. 4. Let $\alpha = (\alpha_A, \alpha_B)$ denote a generic profile of systems of first-order conditional beliefs (which also encodes the plans of Ann & Bob). In the spirit of Dufwenberg & Kirchsteiger (2004) (D&K),¹² define the kindness of Ann toward Bob as the difference between Ann's initial subjective expectation of Bob's monetary payoff and an "equitable" payoff of Bob according to Ann's beliefs:

$$K_{A,B}(\alpha_A) := E[\pi_B, \alpha_A | \emptyset] - \frac{1}{2} \left[\max_{a_A \in \{Out, In\}} E[\pi_B, \alpha_A | a_A] + \min_{a_A \in \{Out, In\}} E[\pi_B, \alpha_A | a_A] \right].$$

Bob's utility is then

$$v_B(z, \alpha) = \pi_B(z) + \theta_B K_{A,B}(\alpha_A) \pi_A(z).$$

Assuming that Ann is selfish, we obtain the p-game in Fig. 7.

As before, we can use Bob's updated second-order beliefs to specify the (expected) utilities he would "experience" at each endnode. If Bob interprets the observed action *In* as evidence that Ann planned to play *In* with probability 1, we obtain the p-game of Fig. 8.

¹²Our account of D&K's approach differs from their own; it is an adaptation that easily fits our framework, which is equivalent to D&K's approach. On belief-dependent reciprocity in games, see also Rabin (1993) and Falk & Fischbacher (2006).

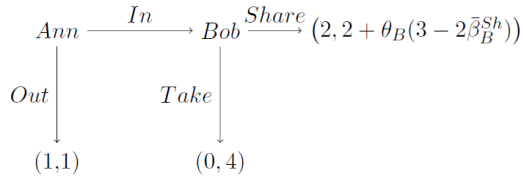


Figure 8: Trust Minigame with exp-utility (reciprocity)

This illustrates how utilities may depend on the game form analyzed. Ann’s action shapes her kindness toward Bob. The degree of kindness depends on the set of alternatives and the material payoff functions (and Ann’s beliefs), hence, on the game form.

2.3 Psychological preferences and time

In standard economic theory, time and agents’ intertemporal preferences are often relevant: present consumption is different from future consumption. Belief-dependent motivation makes time even more important. For instance, when we get frustrated due to an unforeseen loss, we feel angry and our actions are distorted by emotion. However, the time elapsed between the feeling of unexpected loss and the actual decision affects the intensity of anger. This is the so-called cooling-off effect. Other examples concern decay of past emotions, and discounting of future (expected) emotions.

We need to clarify how we model time. In particular, we distinguish between **periods** and **stages**. Periods are the units of time that denote the duration of the interaction among agents: each period is a constant-length interval of time between two successive dates. Importantly, periods affect the psychological utilities of agents through the effects described above. On the other hand, stages denote the moments, within each period, at which agents acquire new information and take actions (which, for simplicity, we also assume are observable). The interpretation is that, in each period, stages unfold in quick succession; hence, they do not affect intertemporal belief-dependent motivations.¹³ Given this, we treat emotions such as anxiety, suspense, frustration, regret, etc. as an “asset” measured at each point in time, stocked in the agent’s mind. That stock, like physical capital, is potentially subject to decay. Anticipated future emotions may be discounted.

To illustrate, consider the Ultimatum Minigame in Fig. 9, explored extensively by BD&S to study frustration and anger. We first analyze a scenario with *one period* and *two stages*. In the first stage, Ann has to make a *Fair* (*Fa*) or a *Greedy* (*Gr*) offer to Bob. In the second stage, if the fair offer was made, then Bob must accept it, else, Bob has to decide whether to *Accept* (*Ac*) or to *Reject* (*Rj*) the greedy offer made by Ann. Monetary payoffs are as seen. We assume that Ann holds risk-neutral standard preferences over monetary payoffs, while Bob holds psychological preferences affected by frustration & anger, as in BD&S. The frustration of Bob, given that Ann made the greedy offer is the difference, if positive, between the amount of money he expected

¹³Such timing of actions and outcomes within a period is—of course—a simplifying assumption.

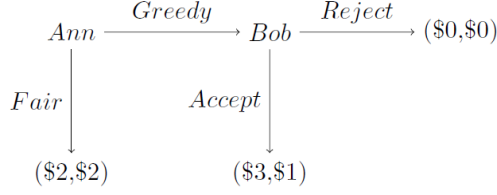


Figure 9: One-period Ultimatum Minigame with material payoffs

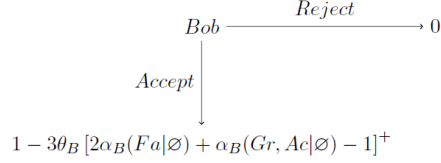


Figure 10: One-period ultimatum Minigame with decision utilities

to get at the root and the maximum amount he can get given the offer. Formally, given Bob's first-order belief α_B , his frustration after *Greedy* is

$$\begin{aligned} F_B[\alpha_B, Gr] &= \max \left\{ 0, E[\pi_B, \alpha_B | \emptyset] - \max_{a_B \in \{Ac, Rj\}} E[\pi_B, \alpha_B | (Gr, a_B)] \right\} \\ &= [2\alpha_B(Fa|\emptyset) + \alpha_B(Gr, Ac|\emptyset) - 1]^+ \end{aligned}$$

where, for every $w \in \mathbb{R}$, $[w]^+ := \max\{0, w\}$. Frustration triggers anger, which here is considered as something that *does not* contribute to experience utility, here identified with expected monetary payoff. In other words, we interpret anger as modeled by BD&S as a distortion affecting the decision utility of actions compared to the expectation of the induced experience utility (in this case, monetary payoff, but, of course in other examples it could be anything, e.g. distributional preferences or guilt aversion). With this, we define the distortion function

$$d_{B, Gr}(z, \alpha) = -\theta_B \pi_A(z) F_B[\alpha_B, Gr]$$

where θ_B is, for simplicity, a commonly known personal trait of Bob measuring the intensity with which his anger distorts his incentives away from material payoff maximization. Thus, the **decision utility** Bob attaches to each reply is the summation of the experienced utility and the expected distortion (see Fig. 10). Bob accepts the greedy offer if

$$\theta_B \cdot 3 [2\alpha_B(Fa|\emptyset) + \alpha_B(Gr, Ac|\emptyset) - 1]^+ < 1.$$

In the second scenario, we consider a game with *two periods* such that each period has only *one stage* (see Fig. 11). Whenever we analyze the interplay between emotions and time we treat

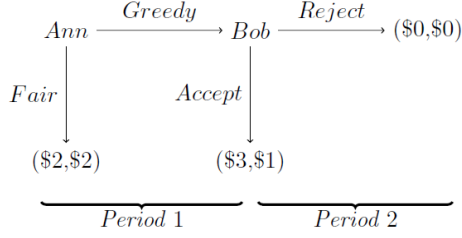


Figure 11: Two-period ultimatum Minigame with material payoffs

the former as a “stock” subject to decay. In this case, the frustration of Bob, right after Ann made the greedy offer is

$$F_B[\alpha_B, Gr] = [2\alpha_B (Fa|\emptyset) + \alpha_B (Gr, Ac|\emptyset) - 1]^+$$

as before. However, in this scenario Bob takes his decision one period later and we consider a cooling-off effect: the distortion function is

$$d_{B,Gr}(z, \alpha) = -\delta_B \theta_B \pi_A(z) F_B[\alpha_B, Gr],$$

where decay factor $\delta_B \in (0, 1)$ is a personal trait of Bob’s. He accepts the greedy offer if

$$\delta_B \theta_B \cdot 3 [2\alpha_B (Fa|\emptyset) + \alpha_B (Gr, Ac|\emptyset) - 1]^+ < 1,$$

which allows for a higher θ_B -threshold compared to the one-period scenario.

3 Basic ingredients: agents, periods, outcomes and beliefs

We now present basic elements of our framework and corresponding mathematical notation. For every set X and $t \in \mathbb{N}$, X^t is the t -fold product of X with generic element $x^t = (x_1, \dots, x_t)$.¹⁴

Individual agents Consider a *finite* index set I of **individual agents** with generic elements i or j . These can be thought of as roles in a game, e.g., buyer or seller. We also consider a fictitious entity called **chance** denoted c .¹⁵ To ease notation, we let $I_c := I \cup \{c\}$, $-i := I_c \setminus \{i\}$ and $-i, c := I_c \setminus \{i, c\}$, that is, $-i$ denotes all individuals different from i , including the pseudo-agent c , whereas $-i, c$ denotes all “real” agents different from i .¹⁶

¹⁴We use standard “blackboard bold” notation for numerical fields, such as the natural numbers \mathbb{N} .

¹⁵In this paper we do not have numerical, or parameterized illustrative examples featuring chance. Yet, given the methodological nature of this contribution, we think it is important to have it in the general framework.

¹⁶We often write $j \neq i$ and $j \neq i, c$ in unions or products of sets indexed by agents to denote respectively the subsets of agents $I_c \setminus \{i\}$ and $I_c \setminus \{i, c\}$.

Profiles Denote functions from I (or I_c) to a generic codomain K with bold letters: $\mathbf{k} \in K^I$, where K^I is the set of functions from I to K . We often consider selections from nonempty-valued correspondences associating each i with a nonempty subset $K_i \subseteq K$, such as the set of types of i . A **profile** of elements of K based on such a correspondence is a $|I|$ -tuple $\mathbf{k} = (k_i)_{i \in I}$, where $k_i \in K_i$ for each $i \in I$. The set of such profiles (functions) is denoted $\mathbf{K} := \prod_{i \in I} K_i \subseteq K^I$. We use similar notation for I_c . When, for $j \in I_c$, K_j is a singleton, we may omit it from our notation, because in this case \mathbf{K} is isomorphic to $\prod_{i \in I \setminus \{j\}} K_i$.

Periods We model time explicitly. Consider a set $\mathbb{T} := \{1, \dots, T\} \subseteq \mathbb{N}$ of time periods, with the interpretation that all the actions take place at the beginning of each period possibly in a multistage fashion, as in, e.g., alternating-offers bargaining models.¹⁷ The generic time period is denoted t , while $T \in \mathbb{N} \cup \{\infty\}$ denotes the horizon (i.e., the maximum duration). If $T = \infty$ (infinite horizon), we have $\mathbb{T} = \mathbb{N}$. A period represents the time between consecutive dates. In particular, the first period starts with date 0 and ends with date 1. Thus, the set of **dates** corresponds to $\mathbb{T}_0 := \{0\} \cup \mathbb{T}$. Whenever $T = 1$, we get an interesting special case, *one period* and *two dates*, which we often consider to simplify aspects of the analysis.

Consequences Consider a space \mathbf{Y} of **collective material outcomes** or **consequences**, e.g. the profiles of agents' monetary payoffs. It is natural to assume a agent-by-agent factorization of the space of material outcomes. For each $i \in I$, Y_i denotes the space of **personal outcomes** for i , and Y_c denotes the space of outcomes for chance. Often, Y_i represents the space of monetary payoffs of i , while elements of Y_c may represent quantities of a public good. We let $\mathbf{Y} := \prod_{i \in I_c} Y_i$ and assume for simplicity that each i observes only $y_i \in Y_i$, as soon as it is realized, and, for every $t \in \mathbb{T}$, i recalls the temporal sequence of his personal outcomes already realized. No agent observes $y_c \in Y_c$. Note that agents may care about the outcomes of other agents.

With more than one period, agents face temporal sequences of outcomes rather than single realizations. If there is a fixed duration T , as in many economic models, an agent's utility may be affected by the whole temporal sequence, or **stream**, of outcomes $\mathbf{y}^T \in \mathbf{Y}^T$.¹⁸ More generally, if the duration is not fixed, we need to consider the set $\mathbf{Y}^{\leq \mathbb{T}} := \bigcup_{t=1}^T \mathbf{Y}^t$ of all (finite and, possibly, infinite) streams of outcomes. For every $t \in \mathbb{T}$, we denote by \mathbf{y}^t (respectively, y_i^t) the stream of the *first* t profiles of realized outcomes (respectively, realized personal outcomes of i).

Exogenous information & personal traits We allow for incomplete information about agents' traits and, possibly, some features of the environment. This is done by letting the outcome of the game and/or agents' utility depend on a profile of parameters comprising exogenously given personal features of each agent, concerning e.g., guilt, disappointment aversion, or altruism, or other characteristics like intelligence or strength. We refer to such features as **personal traits**. The relevant profile of parameters is $(\theta_i)_{i \in I} \in \prod_{i \in I} \Theta_i$, assumed *finite* for simplicity.

¹⁷In the standard alternating-offer bargaining model (e.g., Osborne & Rubinstein, 1994, Ch. 7), each period $t \in \mathbb{N}$ features two stages: offer and response. If the latter is positive, the surplus implied by the agreement just reached is consumed, otherwise—with a one-period delay—play proceeds to the next offer.

¹⁸Note, we implicitly assume that the space of outcomes Y_i of each $i \in I_c$ is time invariant.

The parameter θ_i represents exogenous features *known to i* .¹⁹ There may be also some relevant exogenous features of the environment and/or of the agents that are *unknown to everybody*. For example, some traits of i such as his intelligence may be unknown to i , and i 's beliefs about them may be relevant for the psychological utility of some agent j (possibly i himself, as in the case of self-esteem). We let $\boldsymbol{\theta}_c = (\theta_{c,i})_{i \in I_c} \in \Theta_c$ denote the vector of features unknown to every "real" agent, where $\theta_{c,i}$ ($i \in I$) represents traits of i unknown to i (and everybody else), and $\theta_{c,c}$ represents aspects of the external environment that nobody knows. The set Θ_c (again, *finite*) is called the **space of residual uncertainty**. Thus, $\Theta := \Theta_c \times \prod_{i \in I} \Theta_i$ defines the **space of exogenous uncertainty**.

Beliefs Utilities are characterized by their dependence on agents' beliefs. Following B&D, we allow them to be potentially affected by own beliefs as well as the beliefs of *other* agents. On the one hand, what is truly relevant for agents' preferences are so-called **psychological** or **mental states** (we borrow these terms from Caplin & Leahy 2001) such as anxiety, shame, guilt, anger, etc., felt by agents during their interaction. On the other hand, many of such mental states are triggered by beliefs, including beliefs about beliefs.

Some questions naturally arise: What may agents form beliefs about? Which beliefs are relevant for their utilities? In principle, everything agents may form beliefs about can matter. Each i is uncertain about the profile of parameters $\boldsymbol{\theta}_{-i}$ of exogenous uncertainty concerning others and about the prevailing stream of material outcomes $\mathbf{y}^t \in \mathbf{Y}^{\leq \mathbb{T}}$. Thus, for every $t \in \mathbb{T}_0$, each i forms a belief over $\Theta_{-i} \times \mathbf{Y}^{\leq \mathbb{T}}$, called i 's **space of material uncertainty**. We let $M_i := \Delta(\Theta_{-i} \times \mathbf{Y}^{\leq \mathbb{T}})$ denote the space of i 's beliefs about material uncertainty or **space of material beliefs** of i . For each $i \in I$, we consider streams of material beliefs $(\mu_{i,0}, \mu_i^t) = (\mu_{i,k})_{k=0}^t \in M_i \times M_i^t$ for the first $t + 1$ dates, i.e., for the first t periods. In particular, $\mu_{i,0} \in M_i$ denotes the belief of i at date 0, i.e., his **initial belief**, while each $\mu_{i,t}$ ($t \in \mathbb{T}$) denotes i 's belief at the end of period t . When we consider the simple case of a unique period, a generic stream of material beliefs of i is given by the pair $(\mu_{i,0}, \mu_{i,1}) \in M_i^2$ of initial and terminal material beliefs. When

the duration is not fixed, we need to consider generic streams $(\mu_{i,0}, \mu_i^t) \in M_i^{\leq \mathbb{T}_0} := \bigcup_{t=1}^T M_i \times M_i^t$ of material beliefs of i . For every $t \in \mathbb{T}_0$, denote by $\boldsymbol{\mu}_t \in \mathbf{M} := \prod_{i \in I} M_i$ a generic **profile of**

material beliefs at date t and by $(\boldsymbol{\mu}_0, \boldsymbol{\mu}^t) \in \mathbf{M}^{\leq \mathbb{T}_0} := \bigcup_{t=1}^T \mathbf{M} \times \mathbf{M}^t$ a generic **stream of profiles**

of material beliefs for the first $t + 1$ dates. For each $i \in I$ and $t \in \mathbb{T}$, sets \mathbf{M}_{-i} , \mathbf{M}_{-i}^t and $\mathbf{M}_{-i}^{\leq \mathbb{T}_0}$ are similarly defined. When convenient, we let $M_{i,k}$ denote the set of **date- k material beliefs**. Note that, at each $t \in \mathbb{T}_0$, agent i knows his current material belief $\mu_{i,t}$ and recalls his own previous material beliefs μ_i^{t-1} , but is uncertain about the streams $\boldsymbol{\mu}_{-i}^t \in \mathbf{M}_{-i}^{\leq \mathbb{T}_0}$ of material beliefs of others and about his own future material beliefs $(\mu_{i,k})_{k>t}$, which will depend on yet unobserved information.²⁰

¹⁹More precisely, $i \in I$ knows $\theta_i \in \Theta_i$ and this is common knowledge. In principle, θ_i may contain private information about others' traits, e.g., a teacher may know a pupil's intelligence better than the pupil does.

²⁰This implies that, at date 0, each agent i knows his own initial material belief $\mu_{i,0}$.

4 Psychological utilities without games

We call utilities that are independent of the rules of the specific game forms in which they are embedded “game-independent” (short for “game-form-independent”). They are shaped, in a state-dependent fashion, by the exogenous uncertainty, the (streams of) outcomes, and (streams of) profiles of material beliefs. In Section 6, such utilities are embedded in a game form to obtain a p-game. In particular, we emphasize the role of anticipated and anticipatory feelings.²¹ The date- t stream of outcomes and beliefs $(\mathbf{y}^t, \boldsymbol{\mu}_0, \boldsymbol{\mu}^t)$ determine date- t psychological states that may have negative valence (e.g., disappointment) or positive valence (e.g., surprise). The anticipation of such date- t feelings—that is, **anticipated feelings**—may affect behavior before t , as when we choose a safe option to avoid possible disappointment. In particular, date- t beliefs about *future outcomes* (more generally, about yet unresolved uncertainty) may cause date- t psychological states, called **anticipatory feelings**, with negative valence (e.g., anxiety) or positive valence (e.g., suspense). Thus, the anticipation of such anticipatory feelings, may in turn affect behavior earlier. Mere anticipated feelings, such as disappointment, can be behaviorally relevant with just two dates and one period, whereas the anticipation of anticipatory feelings can be behaviorally relevant with at least three dates and two periods (e.g., the anticipation at date 0 of a possible feeling of anxiety at date 1 concerning a date-2 health outcome). So, we illustrate anticipated feelings in the simple one-period case, and anticipatory feelings in a more complex multi-period case.

One-period case First consider one-period environments ($T = 1$). In this case, the space of streams of outcomes is just \mathbf{Y} and the space of streams of material beliefs is $\mathbf{M}^2 := \prod_{i \in I} (M_{i,0} \times M_{i,1})$ with generic element denoted $(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1)$. For $i \in I_c$, define $X_i := \Theta_i \times Y_i \times M_i^2$. $\mathbf{X} := \prod_{i \in I_c} X_i$.²² In words, each $\mathbf{x} \in \mathbf{X}$ describes all relevant material and mental aspects. Each $x_i \in X_i$ describes i ’s features (i.e., personal traits and beliefs) and outcome, including i ’s initial and terminal beliefs about exogenous uncertainty and collective outcomes. The **one-period psychological utility** of i is defined as $v_i : \mathbf{X} \rightarrow \mathbb{R}$. We can interpret this functional form as a case of *state-dependent utility*, where the state is given by $(\boldsymbol{\theta}, \boldsymbol{\mu}_0, \boldsymbol{\mu}_1)$. This belief-dependent utility makes no reference to the rules of the game. However, as we show in Section 6, it can be easily embedded in any given one-period game form (described in Section 5).

At date 0, each agent i knows θ_i and his initial belief $\mu_{i,0}$, but is uncertain about all the other elements of the utility relevant state \mathbf{x} : The **initial uncertainty space** of i is $\mathbf{X}_{-i} \times M_{i,1}$, where $M_{i,1}$ represents i ’s uncertainty about his own terminal belief, and $\mathbf{X}_{-i} := \Theta_{-i} \times \mathbf{Y} \times \mathbf{M}_{-i}^2$, with generic element denoted by \mathbf{x}_{-i} , represents the uncertainty about other agents and outcomes. Given his known personal trait θ_i and his initial belief $\mu_{i,0}$, agent i forms a belief about $\mathbf{X}_{-i} \times M_{i,1}$ in order to compute the expectation of his experience utility v_i , which we interpret as i ’s “**experienced**” utility.²³

We now present examples of v_i for the one-period case. We assume that, for every $i \in I$, the space of personal outcomes $Y_i \subset \mathbb{R}$ is a space of *monetary payoffs*. The psychological utility of i

²¹See Loewenstein et al. (2001).

²²Notice that, by assumption, X_c is isomorphic to $\Theta_c \times Y_c$, as if chance could hold just one (trivial) belief.

²³Beliefs about \mathbf{X}_{-i} are called “second-order beliefs” because \mathbf{X}_{-i} comprises the space of material beliefs of other agents.

will be determined by means of real-valued functions of the form $f_i : Q \times \Theta_i \rightarrow \mathbb{R}$, where $Q \subseteq \mathbb{R}$ is usually (although not always) a set of “psychological states” of i or some other agent j , such that the section $f_i(\cdot, \theta_i) : \mathbb{R} \rightarrow \mathbb{R}$ is *increasing* and *normalized* to satisfy $f_i(0, \theta_i) = 0$, for each $\theta_i \in \Theta_i$. The interpretation: We first map the stream of beliefs of each agent to psychological states, which—for simplicity—we summarize with real numbers. Function f_i captures how the state q of i or another agent j affects i ’s utility. Common examples of psychological states are disappointment, anxiety or regret. We do *not* assume that all f_i have the same form. The only further restriction we impose is that, whenever Θ_i is an ordered set, section $f_i(q, \cdot)$ is increasing, for every q . Denote generic elements of the class of functions just described by f_i or g_i . For simplicity, one can always assume the following form: $f_i(q, \theta_i) = \theta_i \cdot q$.

Anticipated feelings: disappointment, surprise & guilt In case of disappointment and guilt, outcomes may trigger negative feelings given the initial beliefs of the agents. The *anticipation and expectation* of such negative *feelings* affect behavior. Let Θ_c be a singleton, and Θ_i an ordered set for each i . Agent i is disappointed if his monetary payoff is less than he initially expected: Formally, define the **disappointment operator** of i as

$$\begin{aligned} D_i[\cdot, \cdot] : Y_i \times M_i &\rightarrow \mathbb{R}, \\ (y_i, \mu_i) &\mapsto [\mathbb{E}[\tilde{y}_i, \mu_i] - y_i]^+. \end{aligned}$$

For each pair $(y_i, \mu_{i,0}) \in Y_i \times M_{i,0}$ of i ’s outcome and initial belief, $D_i(y_i, \mu_{i,0})$ measures i ’s disappointment. The anticipation of disappointment affects behavior if D_i enters i ’s utility function, as—for example—in

$$v_i(\boldsymbol{\theta}, \mathbf{y}, \boldsymbol{\mu}_0, \boldsymbol{\mu}_1) = y_i - f_i(D_i[y_i, \mu_{i,0}], \theta_i), \quad (3)$$

where $q \mapsto f_i(q, \theta_i)$ maps disappointment to “utils” according to parameter θ_i .²⁴ Here i ’s utility depends only on his own monetary outcome y_i , trait θ_i , and material belief $\mu_{i,0}$.

Differently, agent i might be (positively) surprised if his monetary payoff is higher than he expected.²⁵ Similarly, we define the **surprise operator** of agent i as

$$\begin{aligned} S_i[\cdot, \cdot] : Y_i \times M_i &\rightarrow \mathbb{R}, \\ (y_i, \mu_i) &\mapsto [y_i - \mathbb{E}[\tilde{y}_i, \mu_i]]^+. \end{aligned}$$

and consider the following formula for i ’s psychological utility:

$$v_i(\boldsymbol{\theta}, \mathbf{y}, \boldsymbol{\mu}_0, \boldsymbol{\mu}_1) = y_i + g_i(S_i[y_i, \mu_{i,0}], \theta_i).$$

We can consider both disappointment and surprise with the formula (cf. Khalmetski et al. 2015):

$$v_i(\boldsymbol{\theta}, \mathbf{y}, \boldsymbol{\mu}_0, \boldsymbol{\mu}_1) = y_i + g_i(S_i[y_i, \mu_{i,0}], \theta_i) - f_i(D_i[y_i, \mu_{i,0}], \theta_i). \quad (4)$$

²⁴In Eq. (3) and in the rest of the article we adopt the following convention: on the left-hand side we write the general expression of the function as presented in its definition, on the right-hand side we write the particular expression of the case under consideration. For example, in Eq. (3), on the left hand side we wrote the general form of the psychological utility of i , which potentially depend on all the elements of \mathbf{x} , while on the right hand side we wrote the specific form assumed by v_i in the example described.

²⁵See Ely et al. (2015) and references therein.

Note that, by definition of f_i and of the operators D_i, S_i , it follows that, for every $\theta_i \in \Theta_i$,

$$g_i(S_i [y_i, \mu_{i,0}], \theta_i) > 0 \implies f_i(D_i [y_i, \mu_{i,0}], \theta_i) = 0$$

and vice versa. As in the general case, we can compute the initial expectation of (4) as

$$E [\tilde{v}_i, \mu_{i,0}] = E [\tilde{y}_i, \mu_{i,0}] + E [g_i(S_i [\cdot, \mu_{i,0}], \theta_i), \mu_{i,0}] - E [f_i(D_i [\cdot, \mu_{i,0}], \theta_i), \mu_{i,0}].$$

Following Battigalli & Dufwenberg (2007), simple guilt can be modeled as the negative feeling of i triggered by the belief that j is disappointed. In two-person games, aversion to guilt can be captured by a utility function of the form:

$$v_i(\boldsymbol{\theta}, \mathbf{y}, \boldsymbol{\mu}_0, \boldsymbol{\mu}_1) = y_i - f_i(D_j [y_j, \mu_{j,0}], \theta_i).$$

Here the state-dependent utility of i depends on the (initial) beliefs of j .

Intrinsic reputation In many social interactions, agents are concerned about the opinion of others and are willing to give up some material payoff (or make a costly effort) in order to improve such opinions, even if this is not expected to yield material benefits in the future. We refer to such concern as “**intrinsic reputation**.”²⁶ Let us consider two-person interactions for simplicity, i.e., $I = \{i, j\}$. The opinion of j about i can be modeled as j 's expectation of i 's “goodness” which, being a personal trait of i , we parametrize through $\theta_i^G \in \Theta_i^G \subseteq \mathbb{R}$. The parameter θ_i^G may measure i 's ability or “goodness of character,” e.g., his altruism or propensity to pro-social behavior. In the first case, θ_i^G will show up in the outcome functions of the game form (see Section 5), as an increase in θ_i^G increases i 's performance. However, we interpret θ_i^G as part of i 's hedonistic preferences, i.e., an argument of v_i . Therefore we can capture this kind of belief-dependent preferences with no reference to the game form. In particular, θ_i^G measures the intensity of i 's other-regarding motivations, such as i 's willingness to increase j 's material payoff.²⁷ Moreover, we parametrize the intensity with which i cares about the opinion of j through $\theta_i^O \in \Theta_i^O \subseteq \mathbb{R}$. Also, θ_i^O is a personal trait, describing the degree of i 's concern for the opinion of others. With this, let $\Theta_i = \Theta_i^G \times \Theta_i^O$ be the parameter space for i 's personal traits. Assuming for simplicity that only the terminal beliefs of j matter to i , we capture the belief-dependent motivations of i with the following utility:

$$v_i(\boldsymbol{\theta}, \mathbf{y}, \boldsymbol{\mu}_0, \boldsymbol{\mu}_1) = y_i + f_i(y_j, \theta_i^G) + g_i\left(E\left[\tilde{\theta}_i^G, \mu_{j,1}\right], \theta_i^O\right). \quad (5)$$

Maps $y_j \mapsto f_i(y_j, \theta_i^G)$ and $E\left[\tilde{\theta}_i^G, \mu_{j,1}\right] \mapsto g_i\left(E\left[\tilde{\theta}_i^G, \mu_{j,1}\right], \theta_i^O\right)$ measure dependence of i 's utility on, respectively, material payoff of j and on j 's estimate of i 's goodness, given $\theta_i = (\theta_i^G, \theta_i^O)$. For example, the prospect of a high (low) value of $E\left[\tilde{\theta}_i^G, \mu_{j,1}\right]$ may induce *pride* (*shame*). Eq. (5)

²⁶See the comments in B&D (Section 6.2) about the dependence of psy-utility on terminal beliefs of others, and the relevant references therein.

²⁷If agent i is also concerned about agent j 's feelings, as in the case of guilt aversion, then also this component may depend on agent j 's beliefs.

generalizes Ellingsen & Johannesson (2008), who consider a linear two-type version.²⁸ Agents' utilities may depend on the personal traits of other agents. In a variation of the previous model, we can replace θ_i^O with the product $\theta_j^G \theta_i^O$ in the g_i function of Eq. (5) so that the more altruistic j is the higher will be the concern of i for the opinion of j ; that is, i would not care about j 's opinion if he knew that j were a bad guy.²⁹

Multiperiod case For the general case with multiple periods ($T \in \mathbb{N} \cup \{\infty\}$), for every $t \in \mathbb{T}$, define $\mathbf{X}^t := \Theta \times \mathbf{Y}^t \times \mathbf{M}_0 \times \mathbf{M}^t$ and $\mathbf{X}^{\leq \mathbb{T}} := \bigcup_{t=1}^T \mathbf{X}^t$. Each $\mathbf{x}^t = (\boldsymbol{\theta}, \mathbf{y}^t, \boldsymbol{\mu}_0, \boldsymbol{\mu}^t) \in \mathbf{X}^{\leq \mathbb{T}}$ describes all the relevant material and mental aspects of our model, taking periods into account. Assuming additive separability and exponential discounting, we posit the following functional form for the **intertemporal psychological utility** of i

$$V_i : \begin{array}{l} \mathbf{X}^{\leq \mathbb{T}} \\ (\boldsymbol{\theta}, \mathbf{y}^t, \boldsymbol{\mu}_0, \boldsymbol{\mu}^t) \end{array} \rightarrow \begin{array}{l} \mathbb{R} \\ \mapsto \sum_{k=1}^t \gamma_i^{k-1} v_i^k(\boldsymbol{\theta}, \mathbf{y}^k, \boldsymbol{\mu}_0, \boldsymbol{\mu}^k), \end{array} \quad (6)$$

where $t \in \mathbb{T}$, $v_i^k : \mathbf{X}^k \rightarrow \mathbb{R}$ is the bounded period- k psy-utility, and the discount factor is such that $\gamma_i \in (0, 1]$ if $T < \infty$ and $\gamma_i \in (0, 1)$ if $T = \infty$. Complete impatience is approximated as $\gamma_i \rightarrow 0$. We assume exponential discounting for the sake of simplicity, because we focus on “exotic” preferences (Loewenstein 2007) due to belief-dependence. As we will show in Section 6, belief-dependence may cause dynamic inconsistency even with exponential discounting. Additional forms of dynamic inconsistency such as present bias can be accounted for by allowing for non-exponential (e.g., quasi-hyperbolic) discounting (Laibson 1997, Phelps & Pollak 1968).

As in the one-period case, agent i is uncertain about all features concerning other agents as well as about his future material outcomes and realized beliefs. Thus, in order to assess the value of his actions he has to consult his belief about such uncertainties, i.e., his second-order belief. We address this in Sections 5 and 6. We next illustrate the interplay of time and psychological utility through some known examples.

Anticipatory feelings: anxiety Caplin & Leahy (2001) argue that interim beliefs about the final outcome may trigger (positive or) negative *anticipatory feelings*. In turn, the anticipation of such feelings may affect earlier behavior. For example, the negative feeling of anxiety can be modeled as an increasing function of the variance of the material payoff. Formally, assume that Θ_c is a singleton, and $\Theta_i = \{0, \bar{\theta}_i\} \subseteq \mathbb{R}$, for each $i \in I$, with $0 < \bar{\theta}_i$, and $Y_i \subseteq \mathbb{R}$ (monetary consequences, or any scalar index of material well-being). On top of this, assume that there is a commonly known fixed duration $T < \infty$, and that all the feasible streams of profiles of outcomes $\mathbf{y}^T \in \mathbf{Y}^T$ have the form $\mathbf{y}^T = (0, \dots, 0, \mathbf{y})$ for some $\mathbf{y} = (y_i)_{i \in I_c}$, that is, all the feasible streams of outcomes deliver a non-zero monetary payoff only in the last period. Thus, we can identify the

²⁸B&D cite several papers on belief-dependent concern for the opinion of others. Furthermore, introspection and experimental evidence suggest that *self-esteem*, i.e., the estimate of own ability, or ability relative to others, affects “ego-utility,” and that the anticipation of such feelings affects behavior (e.g., Kuhnen & Tymula 2012). In this case, v_i depends on i 's updated belief about $\theta_{c,i}$, an exogenous feature of i unknown to i .

²⁹This is somewhat related to the interdependent preference model of Levine (1998), generalized by Gul & Pesendorfer (2016), but these models do not feature belief-dependent utilities.

space of feasible streams of outcomes with \mathbf{Y} , with generic element denoted by \mathbf{y}_T . For every (squared-integrable) function $\varphi : \mathbf{Y} \rightarrow \mathbb{R}$, define the **variance operator** of φ by

$$\begin{aligned} \text{Var} [\varphi, \cdot] : \Delta(\mathbf{Y}) &\rightarrow \mathbb{R}, \\ \mu_i &\mapsto \mathbb{E} \left[(\varphi - \mathbb{E} [\varphi, \mu_i])^2, \mu_i \right]. \end{aligned}$$

We can capture anxiety via the following intertemporal utility of i :

$$V_i(\boldsymbol{\theta}, \mathbf{y}^T, \boldsymbol{\mu}^T) = \gamma_i^T y_{i,T} - \sum_{k=1}^T \gamma_i^{k-1} f_{i,k}(\text{Var} [\cdot, \mu_{i,k}], \theta_i),$$

where, for every $k \in \mathbb{T}$, $f_{i,k} : \mathbb{R}_+ \times \Theta_i \rightarrow \mathbb{R}_+$ is the k -period version of the functions already defined for the one-period case. Note that the non-linearity of i 's psy-utility in his own beliefs typically induces dynamic inconsistencies (cf. Section 7.1).

An agent might also be concerned about the anxiety of others. For example, Caplin & Leahy (2004) analyze a model where a concern for patients' anxiety affects the disclosure policy of doctors. Furthermore, variability of own monetary outcomes (computed with respect to agents' subjective beliefs) do not always cause negative utility: the emotional state of suspense may have positive valence (e.g., Ely et al. 2015 and the references therein).

Section summary We have analyzed game-independent utilities. In our framework, differently from the standard theory of intertemporal decisions, the utility of i potentially depends on the streams of realized beliefs about (external uncertainty and) material outcomes. This may help calibrate agents' utilities by data concerning relatively simple decision situations and then plug them into any game form. Two points are worth noting:

1. First, so far we have treated the streams of realized beliefs of agents as given. However, as it is standard in many applications, such streams of beliefs are the result of processes of belief updating given the acquisition of new information. For example, a rational agent would start the game with a prior belief over his space of uncertainty and would sequentially update that belief according to the new information he acquires through the chain rule of probabilities. Therefore, once an information structure has been defined, the stream of realized beliefs of agent i is pinned down by his initial belief and by the information he sequentially acquires.
2. Second, in many applications, agents' utilities, by their intrinsic nature, cannot be independent of the game form. For example, whenever one aims to model a preference over streams of consequences which takes into account *regret* for past choices then utility will depend on the feasible alternative actions that the agent could have taken and the corresponding material payoffs. Such a model would include in utility regret for past choices through a comparison between realized and (expected) counterfactual outcomes.

The bottom line is clear: several examples of psy-utilities cannot dispense with a game form. We will next describe in detail the rules of the game (Section 5). This is necessary to incorporate

belief-dependent motivations into strategic thinking even with game-independent psy-utilities; for other psy-utilities, like regret avoidance, the description of belief-dependent motivations start directly from the game form (Section 6).

5 Games and Beliefs

In this section, we introduce the rules of the game mathematically described by the game form (Section 5.1), which includes the timing of play, the information flow between stages in each period and between periods, the feasible actions, and outcome functions for each period. Given the information structure, we define conditional probability systems (CPSs), i.e., maps from observable histories to probabilistic beliefs about behavior, and personal traits and beliefs of others (Section 5.2). A state of the game is described by a profile of personal traits θ , a path of play (complete history of actions) z , a profile of first-order CPSs α , and a profile of second-order CPSs β (with β_i consistent with α_i for each i). Each first-order CPS α_i comprises both beliefs about the behavior of others and i 's plan, that is, how i predicts his own behavior. Path z determines a stream of outcomes $(\mathbf{y}_1, \mathbf{y}_2, \dots)$ and—given profile α —a stream of realized beliefs $(\mu_0, \mu_1, \mu_2, \dots)$ about θ and outcomes. These primitive and derived elements will be used in Section 6 to define dynamic p-games.

5.1 Game form

We describe a **game form** with *finite or infinite horizon* and *observable actions*. We begin from a set $\mathbb{T} = \{1, \dots, T\} \subseteq \mathbb{N}$ of periods as in Section 3, and, within each period $t \in \mathbb{T}$, a sequence of $L \in \mathbb{N}$ stages, with L assumed to be fixed for each period. As shown in Section 4, the explicit formalization of the role of time is more important when agents have belief-dependent motivations than in the standard case. Since most textbooks descriptions of games in extensive form do not model time explicitly,³⁰ we do it here with some care. Formally, we add the following concepts to the ones of Section 3.

Stages and actions Each period $t \in \mathbb{T}$ is divided into L stages of arbitrarily short duration, indexed by $\ell \in \{1, \dots, L\}$. Outcomes obtain after the last stage of each period. The sequence of periods represents the passage of time, which affects decay and discounting. Stages, instead, are very short and the unfolding of play in them within a period affects only the flow of information. In each stage, each i takes an action from a (possibly) history-dependent feasible subset of a *finite* set A_i . An **inactive** player is represented as one whose feasible subset consists of just one (pseudo) action denoted by w (for “wait”), assumed to belong to A_i , $i \in I$. Let $\mathbf{A} := \times_{i \in I_c} A_i$ denote the set of **action profiles** $\mathbf{a} = (a_i)_{i \in I_c}$ and, for each $i \in I_c$, let $\mathbf{A}_{-i} := \prod_{j \in I_c \setminus \{i\}} A_j$.

³⁰For example, the general definition of extensive form in Osborne & Rubinstein (1994, Ch.s 6, 11), which is close to ours, does not mention time explicitly. Time enters in specific classes of games, such as bargaining and or repeated games (Ch.s 7, 8) by means of special expressions for the utility of terminal histories.

Histories First, consider generic sequences of action profiles, that is, elements of $\mathbf{A}^{\leq L\mathbb{T}_0} := \bigcup_{\ell=0}^{LT} \mathbf{A}^\ell$, with $\mathbf{A}^0 := \{\emptyset\}$.³¹ We also write $\mathbf{A}^{<L\mathbb{T}_0}$ to denote $\bigcup_{0 \leq \ell < LT} \mathbf{A}^\ell$. We use the standard operation of concatenation of a *finite* sequence with another finite or infinite sequence, from which we define the (weak, hence reflexive) “prefix of” relation \preceq , whose irreflexive part is denoted \prec . Thus, histories are incompletely ordered by \preceq and this makes the set $\mathbf{A}^{\leq L\mathbb{T}_0}$ a **tree** with root \emptyset . Since action profiles are selected in each stage, a history that ends at some stage of period t is a sequence of $(t-1)$ subsequences of length L , possibly followed by a subsequence of length $\ell \leq L$. We let $a_{i,\ell,t}$ denote the action of player i in stage ℓ of period t , and $\mathbf{a}_{\ell,t}$ denotes the action profile selected in stage ℓ of period t . For every $t \in \mathbb{T}$, define:

- incomplete “within-period” histories of length $\ell < L$, $h_t^\ell = (\mathbf{a}_{s,t})_{s=1}^\ell \in \mathbf{A}^\ell$, i.e., h_t^ℓ is the sequence of the first ℓ profiles of actions played in period t ;
- complete “within-period” histories (of length L) $h_t = (\mathbf{a}_{\ell,t})_{\ell=1}^L \in \mathbf{A}^L$, i.e., h_t is the sequence of *all* the profiles of actions played in period t ,³²
- t -period histories of length Lt , $h^t = (h_1, \dots, h_t) \in \mathbf{A}^{Lt}$, i.e., h^t is the sequence of profiles of actions that have been played since the beginning of the game to the end of period t ;
- incomplete $(t+1)$ -period histories of length $Lt + \ell$, $h^{t,\ell} = (h^t, h_{t+1}^\ell) \in \mathbf{A}^{Lt+\ell}$, i.e., $h^{t,\ell}$ is the concatenation of h^t with the first ℓ action profiles played in period $t+1$;
- for infinite-horizon games ($T = \infty$) consider complete infinite histories $h^\infty = (h_t)_{t=1}^\infty$.

The rules of the game determine a subset $\bar{H} \subseteq \mathbf{A}^{\leq L\mathbb{T}_0}$ of feasible sequences of action profiles with the following properties: for all $\bar{h} \in \bar{H}$ and $h \in \mathbf{A}^{\leq L\mathbb{T}_0}$,

- (*closure with respect to prefixes*) if $h \prec \bar{h}$, then $h \in \bar{H}$; hence, the restriction of \preceq to \bar{H} makes \bar{H} a *tree* with root \emptyset ;
- (*period completion*) if the length of \bar{h} is $Lt + \ell$ for some $\ell \in \{1, \dots, L-1\}$, $t \in \mathbb{T}_0$, then $\bar{h} \prec \bar{h}^{t+1}$ for some $\bar{h}^{t+1} \in \bar{H} \cap \mathbf{A}^{L(t+1)}$ (periods are always completed, possibly by taking “waiting” action profiles $\mathbf{w} \in \mathbf{A}$ at each remaining stage of the period);
- (*independent actions*) for each $\bar{h} \in \bar{H}$, the set $\mathcal{A}(\bar{h})$ of feasible action profiles is a Cartesian product: $\mathcal{A}(\bar{h}) = \prod_{i \in I_c} \mathcal{A}_i(\bar{h})$.

We adopt the convention that $\mathcal{A}(\bar{h}) = \emptyset$ if and only if $\mathcal{A}_i(\bar{h}) = \emptyset$ for every $i \in I_c$; $\mathcal{A}(\bar{h}) = \emptyset$ means “game over.” If $T = \infty$ and $\bar{h} \in \bar{H}$ is an infinite history, then $\mathcal{A}(\bar{h}) = \emptyset$ as we do not allow the concatenation of an infinite sequence with another sequence. We let $Z := \{\bar{h} \in \bar{H} : \mathcal{A}(\bar{h}) = \emptyset\}$ denote the set of **terminal histories** with generic element z . Note, the period-completion property implies that each finite-length z is a sequence of complete within-period histories. Also, we let H and \hat{H} respectively denote the sets of **non-terminal** and **finite histories**. Set H

³¹Recall that \emptyset denotes the empty sequence.

³²Thus, the length of each h_t is L . This is a convention, as players may be inactive in some stages.

contains all the (necessarily finite) histories after which players take an action (although the feasible set may be a singleton); thus, $H \subseteq \hat{H}$. Set \hat{H} contains the non-terminal histories and the finite terminal histories. Set \hat{H} corresponds to the set of “observables” events upon which each player conditions his beliefs. Finally, we let $Z(h) := \{z \in Z : h \preceq z\}$ denote the set of terminal histories consistent with $h \in H$.

Outcome functions We now relate players’ personal traits and actions to corresponding streams of material outcomes. As in Section 3, let $\mathbf{Y} = \prod_{i \in I_c} Y_i$ be the set of **collective material outcomes**, i.e., consequences potentially affecting all the players. At the end of the play of period t , an outcome $\mathbf{y} = (y_i)_{i \in I_c} \in \mathbf{Y}$ materializes as a function of the history of actions profiles $h^t \in \mathbf{A}^{Lt}$ and of the parameter profile $\boldsymbol{\theta} \in \Theta$. Formally, for every $t \in \mathbb{T}$, we posit a **period- t outcome function** $\pi_t : \mathbf{A}^{Lt} \times \Theta \rightarrow \mathbf{Y}$. We allow outcomes to depend on $\boldsymbol{\theta}$ because the parameter profile may specify players’ abilities.³³ However, to simplify the analysis, we assume that y_i (e.g., i ’s monetary payoff) depends only on θ_i . Since each i knows θ_i and the parametrized outcome function, and observes past actions, we implicitly assume that each i observes the realized y_i ; but our simplifying assumption rules out the possibility that i can obtain information about the personal traits of others from observing y_i .

Formally, for each $i \in I$ and $t \in \mathbb{T}$, there exists a **personal period- t outcome function** $\pi_{i,t} : \mathbf{A}^{Lt} \times \Theta_i \rightarrow Y_i$ and the period- t collective outcome function has the following form:

$$\begin{aligned} \pi_t : \mathbf{A}^{Lt} \times \Theta &\rightarrow \mathbf{Y}, \\ (\mathbf{a}^{Lt}, \boldsymbol{\theta}) &\mapsto \left(\pi_{c,t}(\mathbf{a}^{Lt}, \boldsymbol{\theta}), (\pi_{i,t}(\mathbf{a}^{Lt}, \theta_i))_{i \in I} \right), \end{aligned}$$

where $y_c = \pi_{c,t}(\mathbf{a}^{Lt}, \boldsymbol{\theta})$ is a component nobody directly observes that may depend on $\boldsymbol{\theta}$.

With this, it is convenient to define, for each $t \in \mathbb{N}$, a **t -cumulative outcome function** that keeps track of the entire stream of collective material outcomes through period t :

$$\begin{aligned} \pi^t : \mathbf{A}^{Lt} \times \Theta &\rightarrow \mathbf{Y}^t, \\ (\mathbf{a}^{Lt}, \boldsymbol{\theta}) &\mapsto \left(\pi_k(\mathbf{a}^{Lk}, \boldsymbol{\theta}) \right)_{k=1}^t. \end{aligned}$$

Similarly, the **t -cumulative outcome function of $i \in I$** is defined by

$$\begin{aligned} \pi_i^t : \mathbf{A}^{Lt} \times \Theta_i &\rightarrow Y_i, \\ (\mathbf{a}^{Lt}, \theta_i) &\mapsto \left(\pi_{i,k}(\mathbf{a}^{Lk}, \theta_i) \right)_{k=1}^t \end{aligned}$$

and records all personal outcomes that i has observed until t . The cumulative outcome function of c is defined similarly, but may depend on $\boldsymbol{\theta}$. By definition, for all $t \in \mathbb{T}$ and $(\mathbf{a}^{Lt}, \boldsymbol{\theta}) \in \mathbf{A}^{Lt} \times \Theta$,

$$\pi^t(\mathbf{a}^{Lt}, \boldsymbol{\theta}) := \left(\pi_c^t(\mathbf{a}^{Lt}, \boldsymbol{\theta}), (\pi_i^t(\mathbf{a}^{Lt}, \theta_i))_{i \in I} \right).$$

Finally, define

$$\begin{aligned} \bar{\pi} : \mathbf{A}^{\leq L\mathbb{T}} \times \Theta &\rightarrow \mathbf{Y}^{\leq \mathbb{T}}, \\ (\mathbf{a}^{Lt}, \boldsymbol{\theta}) &\mapsto \pi^t(\mathbf{a}^{Lt}, \boldsymbol{\theta}) \end{aligned}$$

³³Note that we are not restricting the domain of the outcome function to the feasible histories. This is innocuous and it also makes sense: the rules of the game may specify “input-output” functions that are logically independent of what actions are feasible under different contingencies.

and denote with $\pi = \bar{\pi}|_{Z \times \Theta}$ the restriction of $\bar{\pi}$ to $Z \times \Theta$, i.e., $\pi : Z \times \Theta \rightarrow \mathbf{Y}^{\leq \mathbb{T}}$ is the (θ -dependent) outcome function which assigns a (possibly infinite) stream of outcomes to each (possibly infinite) terminal history $z \in Z$.

5.2 Conditional beliefs

Z-based conditional beliefs We assume that, as a result of strategic reasoning, each $i \in I$ is characterized by a CPS about paths $z \in Z$ (i.e., the *actual* behavior of *everybody*), unknown parameters $\theta_{-i} \in \Theta_{-i}$, and other players’ beliefs about behavior and parameters.³⁴ This approach is different from B&D’s where conditional beliefs were defined over the space of others’ strategies. The most important difference is that, here, we include a representation of each player’s conditional *beliefs about his own behavior*. These beliefs are interpreted as i ’s contingent plan, or strategy. Also, we model conditional beliefs about actual behavior, rather than conditional beliefs about contingent behavior (what a player would do if a history were reached). This deserves discussion.

There are at least two advantages of this Z -based approach. The most important one is conceptual: As in many papers in the epistemic game theory literature, in B&D there is a(n) (explicitly acknowledged) *conflation* between the objective description of how a player would behave at each contingency, which is part of the external (non-mental) state of the world, and the description of how a player plans to choose as a function of what he may observe, which is a mental aspect of the world. This is appropriate if and only if it is assumed to be transparent (i.e., true and always commonly believed) that players carry out their plans. This assumption is restrictive because it rules out the possibility that unexpected moves be interpreted as unintentional mistakes, as in Selten’s (1975) “trembling-hand” story, which is also implicit in K&W’s SE concept. Such conflation can be avoided without adopting the Z -based approach of this paper. One can posit that each player has a CPS about *everybody*’s contingent behavior, including his own.³⁵ Yet, if the external state describes players’ contingent behavior and mental states describe players’ contingent plans, the overall state specifies two versions of each player’s “strategy,” an objective one and a subjective one. This is conceptually legitimate, but may cause confusion. The present approach instead has the pedagogical advantage of featuring, for each player, just one variable that fits the mathematical definition of “strategy”: his subjective plan. There is also a technical advantage of the Z -based approach: the set of pure strategy profiles is exponentially larger than the set of paths.

How do we model players’ beliefs about contingent behavior? The answer is implicit in our heuristic analysis of Section 2: *we model the probability of a conditional event as a conditional probability*. For example, consider the Trust Minigame of Fig. 4. Player i (Ann or Bob) assigns probability p to the conditional “Were Ann to go *In*, Bob would *Share*” if and only if i assigns probability p to *Share* conditional on *In*, i.e., if and only if $\alpha_i(Sh|In) = p$. This is analogous to the connection between mixed and behavior strategies.

We specify players’ conditional beliefs so that we can represent what they would believe upon reaching any finite (terminal or non-terminal) history and how they would assess the expected utility of feasible actions at non-terminal histories. With this in mind, we enrich the set of finite

³⁴See Battigalli et al. (2013), and BD&S.

³⁵See, for example, Battigalli & Siniscalchi (1999) and Battigalli & De Vito (2018).

histories by adding what a player knows after he has just taken an action, but play has not yet moved to the next stage. The set of **personal histories** of $i \in I$ is

$$H_i := \hat{H} \cup \{(h, a_i) \in H \times A_i : a_i \in \mathcal{A}_i(h)\}.$$

Note that we include in H_i also the finite terminal histories because—as explained in Section 4—terminal information may be important for psychological reasons, such as a concern for the opinion of others. We let

$$Z(h_i) := \{z \in Z : \exists \mathbf{a}_{-i} \in \mathcal{A}_{-i}(h), (h, (a_i, \mathbf{a}_{-i})) \preceq z\}$$

denote the set of terminal histories consistent with any personal history $h_i = (h, a_i) \in H_i$. We also define the set $Z(h, \mathbf{a}_{-i})$ in a similar way, for all $h \in H$ and $\mathbf{a}_{-i} \in \mathcal{A}_{-i}(h)$.

As seen in Section 4, each i is uncertain about the streams of profiles of outcomes $\mathbf{y}^t = (y_i^t)_{i \in I_c} \in \mathbf{Y}^{\leq \mathbb{T}}$ and his co-players' traits $\boldsymbol{\theta}_{-i} \in \Theta_{-i}$. Given a game form, terminal history z and parameter $\boldsymbol{\theta}$ uniquely determine the corresponding stream of collective consequences as $\mathbf{y}^{T(z)} = \boldsymbol{\pi}(z, \boldsymbol{\theta})$, where $T(z)$ is the duration implied by z . Therefore, we define the **primitive uncertainty space** for i as the product set $Z \times \Theta_{-i}$ with generic element denoted by the pair $(z, \boldsymbol{\theta}_{-i})$. A **first-order belief** of i is an element of $\Delta(Z \times \Theta_{-i})$. Note that each element of $\Delta(Z \times \Theta_{-i})$ uniquely determines a corresponding belief in $\Delta(\Theta_{-i} \times \mathbf{Y}^{\leq \mathbb{T}})$ implied by outcome functions $\boldsymbol{\pi}$.

To represent players' strategic reasoning in a dynamic game, we specify their beliefs conditional on each personal history. The set of pairs $(z, \boldsymbol{\theta}_{-i})$ consistent with a personal history $h_i \in H_i$ is $Z(h_i) \times \Theta_{-i}$. The latter is the event “ h_i occurred” represented in the primitive uncertainty space $Z \times \Theta_{-i}$. Thus, to ease notation, we let $\alpha_i(E|h_i)$ denote the probability of an event $E \subseteq Z \times \Theta_{-i}$ conditional on $Z(h_i) \times \Theta_{-i}$; similarly, given $h_i \preceq h'_i$, $\alpha_i(h'_i|h_i)$ denotes the probability of $Z(h'_i) \times \Theta_{-i}$ conditional on $Z(h_i) \times \Theta_{-i}$. The set of all maps from H_i to the belief set $\Delta(Z \times \Theta_{-i})$ is $[\Delta(Z \times \Theta_{-i})]^{H_i}$. We consider the subset of such maps that satisfy natural cognitive rationality properties.³⁶

Definition 1 We say that $\alpha_i \in [\Delta(Z \times \Theta_{-i})]^{H_i}$ is a **first-order conditional probability system (CPS)** of i if:

1. **Knowledge implies belief:** for every $h_i \in H_i$, $\alpha_i(h_i|h_i) = 1$;
2. **Chain rule:** For all $h_i, h'_i \in H_i$ and $F \subseteq Z(h'_i) \times \Theta_{-i}$,

$$h_i \preceq h'_i \implies \alpha_i(F|h_i) = \alpha_i(F|h'_i) \alpha_i(h'_i|h_i). \quad (7)$$

3. **Own-action independence (OAI):** for all $h \in H$, $a_i, a'_i \in \mathcal{A}_i(h)$, $\mathbf{a}_{-i} = (a_j)_{j \in I_c \setminus \{i\}} \in \mathcal{A}_{-i}(h)$ and $G \subseteq \Theta_{-i}$,

$$\alpha_i(Z(h, \mathbf{a}_{-i}) \times G|h, a_i) = \alpha_i(Z(h, \mathbf{a}_{-i}) \times G|h, a'_i).$$

³⁶See Battigalli et al. (2013) and BC&S.

Property **1** requires that, upon observing h_i , player i assign probability 1 to $Z(h_i) \times \Theta_{-i}$. Property **2** requires that player i update his beliefs according to the standard rules of conditional probability. Indeed, if $\alpha_i(h'_i|h_i) > 0$ then Eq. (7) can be written as

$$h_i \preceq h'_i \implies \alpha_i(F|h'_i) = \frac{\alpha_i(F|h_i)}{\alpha_i(h'_i|h_i)}.$$

Property **3** requires that i 's beliefs about events that cannot be affected by his action (co-players' traits and simultaneous actions) *do not change* before and after i played his action.

We let $\Delta_{i,1} \subseteq [\Delta(Z \times \Theta_{-i})]^{H_i}$ denote the space of i 's first-order CPSs. With this, $\mathbf{\Delta}_{-i,1} := \prod_{j \in I \setminus \{i\}} \Delta_{j,1}$ and $\mathbf{\Delta}_1 := \prod_{j \in I} \Delta_{j,1}$ respectively denote the spaces of co-players' and all players' profiles of first-order CPS. A triple $(z, \theta, \alpha) \in Z \times \Theta \times \mathbf{\Delta}_1$ completely describes players' actual behavior, their personal traits and their first-order CPSs. Borrowing from Battigalli & Siniscalchi (1999), it can be shown that these spaces have convenient regularity properties (see BC&S for a general treatment).³⁷

Player i is uncertain about $(z, \theta_{-i}, \alpha_{-i})$, i.e., how the game is going to be played, his co-players' personal traits, and what first-order conditional beliefs his co-players would hold upon observing any finite history. We assume, however, that he is certain that the co-players are cognitively rational, i.e., that each α_j ($j \neq i$) satisfies **1-2-3** of Definition 1. Hence, he forms **second-order beliefs** over the space $Z \times \Theta_{-i} \times \mathbf{\Delta}_{-i,1}$ conditional on the occurrence of any personal history $h_i \in H_i$. The occurrence of h_i represented in the first-order uncertainty space $Z \times \Theta_{-i} \times \mathbf{\Delta}_{-i,1}$ corresponds to event $Z(h_i) \times \Theta_{-i} \times \mathbf{\Delta}_{-i,1}$. Let $\beta_i(\cdot|h_i) \in \Delta(Z \times \Theta_{-i} \times \mathbf{\Delta}_{-i,1})$ denote a generic second-order belief of i conditional on $Z(h_i) \times \Theta_{-i} \times \mathbf{\Delta}_{-i,1}$. A system of second-order beliefs of i is denoted $\beta_i = (\beta_i(\cdot|h_i))_{h_i \in H_i}$ and we assume that β_i satisfies cognitive rationality properties similar to **1-2-3** of Definition 1. Such systems of beliefs are called **second-order CPSs**. Let $\Delta_{i,2} \subseteq [\Delta(Z \times \Theta_{-i} \times \mathbf{\Delta}_{-i,1})]^{H_i}$ denote the space of i 's second-order CPSs. Let $\mathbf{\Delta}_{-i,2} := \prod_{j \in I \setminus \{i\}} \Delta_{j,2}$ and $\mathbf{\Delta}_2 := \prod_{j \in I} \Delta_{j,2}$ denote the spaces of profiles of second-order CPSs of, respectively, i 's co-players' and all players. As mentioned in Sections 2 and 4, we do not consider higher-order beliefs, because we assume that *expected* utilities depend only on (first- or) second-order beliefs.

Given a second-order CPS, it is always possible to obtain a first-order CPS by marginalization. In particular, we can define the **marginalization map**

$$\begin{aligned} \text{marg}_i : \Delta_{i,2} &\rightarrow [\Delta(Z \times \Theta_{-i})]^{H_i}, \\ \beta_i &\mapsto \left(\text{marg}_{Z \times \Theta_{-i}} \beta_i(\cdot|h_i) \right)_{h_i \in H_i}. \end{aligned}$$

For each second-order CPS β_i , $\text{marg}_i(\beta_i)$ corresponds to the first-order CPS obtained by marginalization of $\beta_i(\cdot|h_i)$ for every $h_i \in H_i$. One can show that $\text{marg}_i(\Delta_{i,2}) = \Delta_{i,1}$, i.e., for every $\beta_i \in \Delta_{i,2}$, $\left(\text{marg}_{Z \times \Theta_{-i}} \beta_i(\cdot|h_i) \right)_{h_i \in H_i}$ is a first-order CPS.

Realized beliefs Strategic reasoning affects what any given player i would believe conditional on each personal history, i.e., his CPS. We assume, however, that the utility attached by i to any

³⁷In particular, for each $i \in I$, $Z \times \Theta_{-i}$ is compact, metrizable and each set $Z(h_i) \times \Theta_{-i}$ ($h_i \in H_i$) is clopen (closed & open). It follows that $\Delta_{i,1}$, $\mathbf{\Delta}_{-i,1}$ ($i \in I$), and $Z \times \Theta \times \mathbf{\Delta}_1$ are compact metrizable.

terminal history z depends only on realized beliefs. Here we derive the realized beliefs implied by a player's first-order CPS and a terminal history. Fix $i \in I$. Every pair $(\alpha_i, z) \in \Delta_{i,1} \times Z$ yields a stream of realized first-order beliefs

$$(\alpha_i(\cdot|h^t(z)))_{t=0}^{T(z)} \in \bigcup_{t \in \mathbb{T}} [\Delta(Z \times \Theta_{-i})]^t,$$

where, $T(z) \in \mathbb{T}$ is the (possibly infinite) duration implied by z , $h^t(z)$ is the unique complete t -period history weakly preceding z , and $h^0(z) = \emptyset$. This in turn—for every $\theta_i \in \Theta_i$ —yields a stream of material beliefs $(\mu_{i,t})_{t=0}^{T(z)}$, i.e., beliefs about streams of material outcomes and co-players' traits. To see this in detail, recall that $\pi : Z \times \Theta \rightarrow \mathbf{Y}^{\leq \mathbb{T}}$ is the outcome function that associates each pair (z, θ) with the corresponding stream of outcomes. For each measurable set $F \subseteq \Theta_{-i} \times \mathbf{Y}^{\leq \mathbb{T}}$, the probability of F implied by (z, θ_i, α_i) is

$$\psi_{i,t}(z, \theta_i, \alpha_i)(F) := \alpha_i(\{(z', \theta'_{-i}) \in Z \times \Theta_{-i} : (\theta'_{-i}, \pi(z', \theta_i, \theta'_{-i})) \in F\} | h^t(z)).$$

Thus, we obtain a t -period realized (or on-path) material-belief function $\psi_{i,t} : Z \times \Theta_i \times \Delta_{i,1} \rightarrow M_i$, and we define the **realized-belief function** of i as

$$\begin{aligned} \psi_i : Z \times \Theta_i \times \Delta_{i,1} &\rightarrow M_i^{\leq \mathbb{T}_0}, \\ (z, \theta_i, \alpha_i) &\mapsto (\psi_{i,t}(z, \theta_i, \alpha_i))_{t=0}^{T(z)}. \end{aligned}$$

Let $\boldsymbol{\psi} := (\psi_i)_{i \in I}$. Thus, $(\boldsymbol{\theta}, \boldsymbol{\pi}(z, \boldsymbol{\theta}), \boldsymbol{\psi}(z, \boldsymbol{\theta}, \boldsymbol{\alpha})) \in \mathbf{X}^{\leq \mathbb{T}}$ is the stream of realized (on path) outcomes and beliefs induced by $(z, \boldsymbol{\theta}, \boldsymbol{\alpha})$.

Belief factorization For every $i \in I$, define $\Sigma_i := \prod_{h \in H} \Delta(\mathcal{A}_i(h))$ and $\Sigma_{-i} := \prod_{h \in H} \Delta(\mathcal{A}_{-i}(h))$.

Technically, each $\sigma_i \in \Sigma_i$ is a **behavior strategy**, whereas $\boldsymbol{\sigma}_{-i} \in \Sigma_{-i}$ is a kind of “correlated” behavior strategy of the co-players (including c) that we interpret as a conjecture of i about others' history-dependent behavior. Recalling that, $\alpha_i(Z(h, a_i) \times \Theta_{-i}|h)$ is the conditional probability of action a_i given h ($\alpha_i(Z(h, \mathbf{a}_{-i}) \times \Theta_{-i}|h)$ has a similar meaning), we obtain i 's plan and his conjecture about co-players' behavior as follows:

$$\begin{aligned} \hat{\sigma}_i : \Delta_{i,1} &\rightarrow \Sigma_i, \\ \alpha_i &\mapsto \left((\alpha_i(Z(h, a_i) \times \Theta_{-i}|h))_{a_i \in \mathcal{A}_i(h)} \right)_{h \in H}, \\ \hat{\boldsymbol{\sigma}}_{-i} : \Delta_{i,1} &\rightarrow \Sigma_{-i}, \\ \alpha_i &\mapsto \left((\alpha_i(Z(h, \mathbf{a}_{-i}) \times \Theta_{-i}|h))_{\mathbf{a}_{-i} \in \mathcal{A}_{-i}(h)} \right)_{h \in H}. \end{aligned}$$

OAI implies that, for all $h \in H$, $a_i \in \mathcal{A}_i(h)$, and $\mathbf{a}_{-i} \in \mathcal{A}_{-i}(h)$,

$$\hat{\boldsymbol{\sigma}}_{-i}(\alpha_i)(\mathbf{a}_{-i}|h) = \alpha_i(Z(h, \mathbf{a}_{-i}) \times \Theta_{-i}|h, a_i),$$

which yields the following natural factorization: for every $\mathbf{a} \in \mathcal{A}(h)$,

$$\hat{\sigma}_{i,I}(\alpha_i)(\mathbf{a}|h) := \alpha_i(Z(h, \mathbf{a}) \times \Theta_{-i}|h) = \hat{\sigma}_i(\alpha_i)(a_i|h) \times \hat{\boldsymbol{\sigma}}_{-i}(\alpha_i)(\mathbf{a}_{-i}|h).$$

We have completed the description of the primitive and main derived elements of our analysis. Next, we will put them together to incorporate belief-dependent motivations into game forms to obtain p-games.

6 Dynamic p-games

In previous sections, we presented game-independent psychological utilities, multiperiod game forms, and systems of conditional beliefs based on such game forms. Here we use these ingredients to obtain the full specification of a dynamic p-game. We focus on **first-order p-games**, i.e., the utility of each i potentially depends only on i 's and his co-players' first-order beliefs.³⁸ We start with the simple case of one-period multistage p-games ($T = 1$ according to the notation of Section 5) and then move on to multiperiod games ($T \in \mathbb{N} \cup \{\infty\}$). To obtain a p-game, we either embed game-independent psychological utility functions (Section 4) in a game form (Section 5), or—*for game dependent preferences*—directly define on the game form the belief-dependent “experience utility” of a terminal history and “decision utility” of actions.

6.1 One-period p-games

Suppose that there is only one period. Within this period there are L stages; thus, if $L > 1$, players make sequential choices. The set of terminal histories comprises sequences of action profiles of length L : $Z \subseteq \mathbf{A}^L$.³⁹ Moreover, outcomes (consequences) materialize at the end of the game, and streams of beliefs consist of an initial and a terminal belief for each player.

Belief-dependent utility of paths Start with a game-independent “experience” utility function as in Section 4. In the one-period case, the utility of i has the form $v_i : \Theta \times \mathbf{Y} \times \mathbf{M}^2 \rightarrow \mathbb{R}$, where $\mathbf{M} = \prod_{i \in I} \Delta(\Theta_{-i} \times \mathbf{Y})$. Section 5 shows how to map terminal histories and first-order conditional beliefs (CPSs) of players to the corresponding outcomes and realized beliefs by means of the outcome functions $\pi = (\pi_i : Z \times \Theta_i \rightarrow Y_i)_{i \in I_c}$ and realized-belief functions $\psi = (\psi_i : Z \times \Theta_i \times \Delta_{i,1} \rightarrow M_i^2)_{i \in I}$. Plugging π and ψ into v_i we obtain a reduced-form function that gives the parameterized belief-dependent utility of paths:

$$\begin{aligned} u_i : Z \times \Theta \times \Delta_1 &\rightarrow \mathbb{R}, \\ (z, \theta, \alpha) &\mapsto v_i(\theta, \pi(z, \theta), \psi(z, \theta, \alpha)). \end{aligned} \tag{8}$$

We illustrate Eq. (8) with reference-dependence à la Koszegi & Rabin (2006).

Reference-dependent utility Fix any player i and let Y_i be a closed interval in \mathbb{R}_+ , with the interpretation that $y_i \in Y_i$ is i 's consumption. Assuming that θ is common knowledge, and given that we are not making comparative static exercises, we need not make the dependence of utilities or outcomes on θ explicit. Let

$$v_i(\mathbf{y}, \boldsymbol{\mu}_0, \boldsymbol{\mu}_1) = y_i + f_i(y_i - \mathbb{E}[\tilde{y}_i, \mu_{i,0}])$$

where the increasing function f_i captures how much i cares about the difference between his actual consumption y_i and the reference point $\mathbb{E}[\tilde{y}_i, \mu_{i,0}]$, his initially expected consumption.

³⁸For a general treatment of k -th order psychological games, see BC&S.

³⁹If the actual game being represented has also shorter terminal histories, they appear in our notation with players being forced to “wait” in later stages of the game.

Given a one-period game form with the corresponding outcome and realized beliefs functions, we obtain the belief-dependent utility of paths

$$\begin{aligned} u_i(z, \boldsymbol{\alpha}) &= \pi_i(z) + f_i(\pi_i(z) - \mathbb{E}[\pi_i, \psi_{i,0}(z, \boldsymbol{\alpha}_i)]) \\ &= \pi_i(z) + f_i(\pi_i(z) - \mathbb{E}[\pi_i, \boldsymbol{\alpha}_i | \emptyset]). \end{aligned}$$

In game-dependent cases, we define $u_i : Z \times \Theta \times \Delta_1 \rightarrow \mathbb{R}$ directly on the game form to model the relevant belief-dependent motivation as we did in Section 2 for reciprocity in the Trust Minigame. From now on, we take the profile of game-based utility functions $(u_i)_{i \in I}$ as a primitive of the analysis. These functions are similar to those in B&D. The most important differences are that (i) B&D allow for dependence on higher-order beliefs, and (ii) B&D do not consider beliefs about own behavior (see related comments in Section 1).

Decision utility, experienced utility, distortion, and action tendency Consider player i at the beginning of stage ℓ , hence, after he observed a nonterminal history $h = (\mathbf{a}_s)_{s=1}^{\ell-1}$ of length $\ell-1$. His goal is to choose an action with the highest expected utility. Since u_i (typically) depends on unknown co-players' beliefs, i consults his conditional second-order beliefs to determine the expected utility of actions. We first derive the expectation of i 's experience utility given h and his action. Then we argue that his actual action tendencies may be described by the maximization of a different “decision utility” obtained by “distorting” the expectation of u_i .

The expected experience utility of action $a_i \in \mathcal{A}_i(h)$ given h and personal features (θ_i, β_i) (expressed as a Lebesgue integral) is

$$\mathbb{E}[u_i(\cdot, \theta_i, \alpha_i), \beta_i | h, a_i] = \int_{Z \times \Theta_{-i} \times \Delta_{-i,1}} u_i(z, \theta_i, \boldsymbol{\theta}_{-i}, \alpha_i, \boldsymbol{\alpha}_{-i}) \beta_i(dz, d\boldsymbol{\theta}_{-i}, d\boldsymbol{\alpha}_{-i} | h)$$

(with $\alpha_i = \text{marg}_i(\beta_i)$).

In some applications, the choice of player i at history h is also driven by some “local” effects, like the urge to harm others caused by frustration at h due to perceived goal obstruction (BD&S), the desire to repair harm caused by a guilt-eliciting event, to insulate oneself from negative evaluation when ashamed (Tangney 1995), or to reciprocate (un)kind behavior (Dufwenberg & Kirchsteiger 2004). We abstractly model such effects by means of **distortion functions** $(d_{i,h} : Z \times \Theta \times \Delta_1 \rightarrow \mathbb{R})_{h \in H}$. The expectation of $d_{i,h}$ given h , a_i , and (θ_i, β_i) is

$$\mathbb{E}[d_{i,h}(\cdot, \theta_i, \alpha_i), \beta_i | h, a_i] = \int_{Z \times \Theta_{-i} \times \Delta_{-i,1}} d_{i,h}(z, \theta_i, \boldsymbol{\theta}_{-i}, \alpha_i, \boldsymbol{\alpha}_{-i}) \beta_i(dz, d\boldsymbol{\theta}_{-i}, d\boldsymbol{\alpha}_{-i} | h)$$

(with $\alpha_i = \text{marg}_i(\beta_i)$).

With this, we obtain the “local” **decision utility functions** $(\bar{u}_{i,h} : \mathcal{A}_i(h) \times \Theta_i \times \Delta_{i,2} \rightarrow \mathbb{R})_{h \in H}$ such that, for all $h \in H$, $a_i \in \mathcal{A}_i(h)$, $\theta_i \in \Theta_i$, $\beta_i \in \Delta_{i,2}$,

$$\bar{u}_{i,h}(a_i, \theta_i, \beta_i) = \mathbb{E}[u_i(\cdot, \theta_i, \alpha_i), \beta_i | h, a_i] + \mathbb{E}[d_{i,h}(\cdot, \theta_i, \alpha_i), \beta_i | h, a_i].$$

We assume that i at h maximizes his decision utility $\bar{u}_{i,h}(\cdot, \theta_i, \beta_i)$. Thus, we define, for each $h \in H$, the best reply correspondence

$$\begin{aligned} r_{i,h} : \quad \Theta_i \times \Delta_{i,2} &\rightrightarrows \mathcal{A}_i(h), \\ (\theta_i, \beta_i) &\mapsto \arg \max_{a_i \in \mathcal{A}_i(h)} \bar{u}_{i,h}(a_i; \theta_i, \beta_i). \end{aligned}$$

Since $\mathcal{A}_i(h)$ is assumed to be finite, $r_{i,h}(\theta_i, \beta_i)$ is nonempty for all θ_i and β_i .

To sum up, a one-period **multistage p-game** is a structure

$$\left\langle I, A_c, \Theta_c, \bar{H}, (A_i, \Theta_i, u_i, (d_{i,h})_{h \in H})_{i \in I} \right\rangle,$$

where $\bar{H} \subseteq \mathbf{A}^L$ satisfies the properties described in Section 5, and, for all $i \in I$ and $h \in H$, the functions u_i and $d_{i,h}$ are respectively the experience utility function and the local distortion function (at h) of i .

In some interesting applications preferences in p-games can be represented by “almost standard” expected utility formulas. Suppose that there are *no distortions* ($d_{i,h} \equiv 0$ for all i and h) and that the experience utility function u_i is derived from a game-independent one v_i , as per Eq. (8), that does not depend on i ’s own beliefs about material outcomes, as in models of simple guilt aversion and intrinsic reputational concerns (but unlike models of expectation-based reference dependence, or anticipatory feelings, see Section 4). Then player i maximizes the expected value of u_i , which does not depend on his beliefs α_i . In this case u_i looks just like a classical state-dependent utility, that is, a utility that depends on paths (or outcomes), parameters known to i (θ_i and α_i), and some parameters or variables unknown to i (θ_{-i} and α_{-i}). This implies that all the standard properties of subjective expected utility hold. In particular, preferences are *dynamically consistent* and *randomization is superfluous*, i.e., optimal planning can be restricted to pure strategies. The same conclusion can be reached under a weaker assumption. Since i ’s beliefs about co-players are a known given for his planning process, u_i may depend on such beliefs; the key assumption (besides no distortions) is that u_i does not depend on i ’s plan. Let $\alpha_{i,-i}$ denote i ’s system of conditional beliefs about others.⁴⁰ If α_i and α'_i are such that $\alpha_i \neq \alpha'_i$, but $\alpha_{i,-i} = \alpha'_{i,-i}$, then the only relevant difference between these two CPSs is that they yield different plans for i , that is, $\hat{\sigma}_i(\alpha_i) \neq \hat{\sigma}_i(\alpha'_i)$. If i ’s plan does not affect his preferences, then his experience utility is unaffected by such change.

Definition 2 *The preferences of player i satisfy **own-plan independence** (OPI) if there are no distortions and, for all $z \in Z$, $\theta \in \Theta$, $\alpha_i, \alpha'_i \in \Delta_{i,1}$, $\alpha_{-i} \in \Delta_{-i,1}$,*

$$\alpha_{i,-i} = \alpha'_{i,-i} \implies u_i(z, \theta, \alpha_i, \alpha_{-i}) = u_i(z, \theta, \alpha'_i, \alpha_{-i}).$$

6.2 Examples of game-dependent psychological utilities (one period)

Some belief-dependent preferences are derived within the context of a specific causal structure captured by the game tree and the outcome functions. Regret and anger are cases in point. We now present some examples which show how to model these aspects through game-form-dependent psychological preferences. For the sake of simplicity, we consider monetary outcomes ($\mathbf{Y} \subseteq \mathbb{R}^{I_c}$) and let the outcome functions depend only on the sequence (of length L) of actions chosen by players, that is, $\pi_{i,1} : \mathbf{A}^L \rightarrow Y_i$, for every $i \in I_c$.⁴¹

⁴⁰Formally, $\alpha_{i,-i} = (\alpha_i(Z(\mathbf{a}_{-i}, h) \times \{\theta_{-i}\} | h))_{h \in H, \theta_{-i} \in \Theta_{-i}}$.

⁴¹Despite this, the utility of terminal histories may depend on θ .

One-period regret Assume that player i primarily cares about his monetary outcome $y_i \in Y_i$ and that he experiences regret at the end of the game if he obtains material payoff $y_i = \pi_i(z)$ and believes that he could have done better. For every $(z, \alpha_i) \in Z \times \Delta_{i,1}$ define the ex post belief $\sigma_{i,-i}(z, \alpha_i) \in \Sigma_{-i}$ of i as

$$\sigma_{i,-i}(z, \alpha_i)(\mathbf{a}_{-i}|h) := \begin{cases} 1 & \text{if } (h, \mathbf{a}_{-i}) \preceq z \\ \hat{\sigma}_{i,-i}(\alpha_i)(\mathbf{a}_{-i}|h) & \text{else} \end{cases}$$

for all $h \in H$ and $\mathbf{a}_{-i} \in \mathcal{A}_{-i}(h)$. We interpret $\sigma_{i,-i}(z, \alpha_i)$ as the ex post belief of i concerning his co-players' contingent behavior given that he observed z and that his first-order belief is α_i . Also, we define the function $\zeta : \Sigma_i \times \Sigma_{-i} \rightarrow \Delta(Z)$ that maps plans of i and his beliefs about his co-players' behavior to subjective probability distributions over terminal paths. With this, the regret of i given (z, α_i) is defined as

$$\text{Re}_i[z, \alpha_i] := \max_{\sigma_i \in \Sigma_i} \{E[\pi_{i,1}, \zeta(\sigma_i, \sigma_{i,-i}(z, \alpha_i))] - \pi_{i,1}(z)\}.$$

In words, the regret of player i at terminal history z given α_i is equal to the difference between the maximum monetary payoff he could have reached given the ex post belief $\sigma_{i,-i}(z, \alpha_i)$ and the actual monetary payoff he received, i.e., $\pi_{i,1}(z)$. In particular, the maximum is taken over all the feasible plans $\sigma_i \in \Sigma_i$ of player i .⁴² By definition, regret is non-negative.

The psychological utility of i at $(z, \boldsymbol{\theta}, \boldsymbol{\alpha})$ is

$$u_i(z, \boldsymbol{\theta}, \boldsymbol{\alpha}) = \pi_{i,1}(z) - f_i(\text{Re}_i[z, \alpha_i], \theta_i).$$

Note that the definition of Re_i cannot dispense with the given game form. Indeed, both the outcome function of i and the feasible actions of players play a fundamental role establishing the extent of the regret of i . Finally, for all pairs $(\theta_i, \beta_i) \in \Theta_i \times \Delta_{i,2}$, nonterminal histories $h \in H$, and actions $a_i \in \mathcal{A}_i(h)$, we obtain that the decision utility of a_i given (θ_i, β_i) at h :

$$\bar{u}_{i,h}(a_i; \theta_i, \beta_i) = E[\pi_{i,1}, \alpha_i | h, a_i] - E[f_i(\text{Re}_i[\cdot, \alpha_i], \theta_i), \alpha_i | h, a_i]$$

(with $\alpha_i = \text{marg}_i(\beta_i)$). Note that we assumed no distortion here, and that—by definition—regret does not depend on $\hat{\sigma}_i(\alpha_i)$. Therefore u_i satisfies *own-plan independence*.

It also makes sense to consider players that are concerned with others' regret (e.g., parents may dislike regret of their children). For example, let $\Theta_i = \prod_{j \in I} \Theta_{i,j}$, where, for every $j \in I$, $\theta_{i,j} \in \Theta_{i,j}$ represents the intensity of i 's concern for j 's regret. In this case u_i is defined as

$$u_i(z, \boldsymbol{\theta}, \boldsymbol{\alpha}) = \pi_{i,1}(z) - \sum_{j \in I} f_i(\text{Re}_j[z, \alpha_j], \theta_{i,j})$$

and the resulting decision utility at any $h \in H$ is defined by

$$\bar{u}_{i,h}(a_i; \theta_i, \beta_i) = E[\pi_{i,1}, \alpha_i | h, a_i] - \sum_{j \in I} E[f_i(\text{Re}_j[\cdot], \theta_{i,j}), \beta_j | h, a_i]$$

(with $\alpha_i = \text{marg}_i(\beta_i)$). This is a belief-dependent form of other-regarding preferences and decision utility crucially depends on second-order beliefs.

⁴²By standard arguments, for every maximizer σ_i^* , every deterministic (pure) plan in the “support” of σ_i^* is also a maximizer.

One-period frustration & anger Following BD&S, we assume for simplicity that the experience utility of each player i coincides with his monetary payoff: $u_i(z, \theta, \alpha) = \pi_{i,1}(z)$. The frustration of i at any nonterminal history $h \in H$ given α_i is the gap, if positive, between the payoff i expected at the beginning of the game and the maximum achievable expected payoff given h :

$$F_i[h, \alpha_i] := \left[\mathbb{E}[\pi_{i,1}, \alpha_i | \emptyset] - \max_{a_i \in \mathcal{A}_i(h)} \mathbb{E}[\pi_{i,1}, \alpha_i | h, a_i] \right]^+ . \quad (9)$$

Note that in the first stage, hence conditional on the empty history $h = \emptyset$, frustration $F_i[h, \alpha_i]$ is null. Indeed, Eq. (9) implies that negative surprise is a necessary (although not sufficient) condition for frustration, and players cannot be surprised at the beginning of the game. Thus, frustration can affect behavior only in games with at least two stages. Next, we define the **simple anger** distortion function of i at $h \in H$ as

$$d_{i,h}(z, \theta, \alpha) = -\theta_i F_i[h, \alpha_i] \sum_{j \in I \setminus \{i\}} \pi_{j,1}(z)$$

for every $(z, \theta, \alpha) \in Z \times \Theta \times \Delta_1$. With this, the decision utility of action $a_i \in \mathcal{A}_i(h)$ at $h \in H$, given (θ_i, β_i) is defined by

$$\bar{u}_{i,h}(a_i; \theta_i, \beta_i) = \mathbb{E}[\pi_{i,1}, \alpha_i | h, a_i] + \mathbb{E}[d_{i,h}, \alpha_i | h, a_i] .$$

The action tendency of players affected by simple anger is to increase the aggregate harm that can be inflicted on co-players, if not too costly. BD&S consider more nuanced versions of anger where aggression is only directed to blameworthy co-players.⁴³ With more than two stages, this model corresponds to the “fast play” version of the multistage setting in BD&S, interpretable as a one-period game. The key issue is that, in one-period games, the beliefs that determine the reference point are those held at the beginning of the game.

6.3 Multiperiod p-games

Now we extend the analysis to an arbitrary horizon $T \in \mathbb{N} \cup \{\infty\}$. In Section 4 we put forward a representation of game-independent preferences over streams of outcomes and realized beliefs by an additive discounted aggregation of a stream of one-period utilities. Given a multi-period game form, this yields belief-dependent intertemporal (experience) utilities of paths:

$$u_i : \begin{array}{l} Z \times \Theta \times \Delta_1 \rightarrow \mathbb{R}, \\ (z, \theta, \alpha) \mapsto \sum_{t=1}^{T(z)} \gamma_i^{t-1} v_i^t(\theta, \pi^t(z, \theta), \psi^t(z, \theta, \alpha)), \end{array} \quad (10)$$

where $\gamma_i \in (0, 1)$ is the discount factor,⁴⁴ and the (π^t, ψ^t) functions give the t -period streams of material outcomes and beliefs.⁴⁵

⁴³In leader-follower games, simple anger is equivalent to *anger from blaming behavior*. See BD&S.

⁴⁴If $T < \infty$ we may allow $\gamma_i = 1$ (no discounting).

⁴⁵See Section 5. To ease notation, we write $\pi^t(z, \theta)$ instead of $\pi^t(h^t(z), \theta)$. We also let $(\mu_0, \mu^t) = \psi^t(z, \theta, \alpha)$ denote the on-path beliefs of the first t periods, hence the first $t + 1$ dates.

As in the one-period case, game-dependent psychological motivations can be captured by the reduced-form experience utility of Eq. (10), which is similar to the (first-order) belief-dependent utility of B&D. We now take the profile of utility functions $(u_i)_{i \in I}$ as given. Similarly, for the sake of simplicity, we do not specify a particular form (e.g., additively separable) for the distortion functions and just posit a profile $(d_{i,h} : Z \times \Theta \times \Delta_1 \rightarrow \mathbb{R})_{h \in H}$. From this, as before we obtain, for each player $i \in I$, the decision utility functions $(\bar{u}_{i,h} : \mathcal{A}_i(h) \times \Theta_i \times \Delta_{i,2} \rightarrow \mathbb{R})_{h \in H}$ which yield “local” best reply correspondences $(r_{i,h} : \Theta_i \times \Delta_{i,2} \rightrightarrows \mathcal{A}_i(h))_{h \in I}$ as in the one-period case.

The models of, e.g., anxiety avoidance (Caplin & Leahy 2001), suspense and surprise (Ely et al. 2015), and guilt aversion in a repeated game (Attanasi et al. 2018) fit this setup. The following example provides an illustration with game-dependent preferences and also allows for a comparison between multiperiod games and one-period multistage games.

Multiperiod frustration and anger First, consider the simpler case with a unique stage for every period (i.e., $L = 1$). In this case, *stages and periods coincide*. Similarly to the one-period case, suppose that one-period experience utility coincides with the monetary payoff. Fix $t \in \mathbb{T}$, and consider histories $h^{t-1}, h^t \in H$ with $h^{t-1} \prec h^t$.⁴⁶ Recall that *there can be no frustration in the first stage* (hence, period). Building on the one-period case, we model the *flow* of frustration experienced in stage $t + 1$, as the gap, if positive, between the present value of payoffs expected at the beginning of stage t —hence, conditional on h^{t-1} —and the maximum expected present value achievable in stage $t + 1$, given h^t . Formally, let $\Pi_{i,t}(z) = \sum_{k \geq t} \gamma_i^{k-t} \pi_{i,k}(h^k(z))$ denote the present value of payoffs from t given path z , then define

$$F_{i,t+1}[h^t, \alpha_i] := \left[\mathbb{E}[\Pi_{i,t}, \alpha_i | h^{t-1}] - \max_{a_i \in \mathcal{A}_i(h^t)} \mathbb{E}[\Pi_{i,t}, \alpha_i | h^t, a_i] \right]^+. \quad (11)$$

The flow of frustration experienced in stage $t+1$ contributes to the *stock* of frustration experienced in the previous t stages, but the effect of frustration in early stages fades exponentially according to decay factor $\delta_i \in [0, 1)$ (cooling-off effect). Formally, the **stock** of frustration experienced at the beginning of stage $t + 1$ is

$$F_i^{t+1}[h^t, \alpha_i] := F_{i,t+1}[h^t, \alpha_i] + \sum_{k=2}^t \delta_i^{t-k+1} F_{i,k}[h^{k-1}, \alpha_i],$$

where, for each $k \leq t - 1$, h^k denotes the unique prefix of h^t of length k . The stock of frustration cumulated in the first t stages affects the decision of i in $t + 1$ given h^t through the (simple anger) distortion function

$$d_{i,h^t}(z, \theta, \alpha) := -\theta_i F_i^{t+1}[h^t, \alpha_i] \sum_{j \in I \setminus \{i\}} \Pi_{j,t+1}(z).$$

With this, the decision utility given h^t of any action $a_i \in \mathcal{A}_i(h^t)$ is

$$\bar{u}_{i,h^t}(a_i; \theta_i, \beta_i) := \mathbb{E}[\Pi_{i,t+1}, \alpha_i | h^t, a_i] + \mathbb{E}[d_{i,h^t}, \alpha_i | h^t, a_i].$$

⁴⁶Recall that, if $t = 1$, then $h^{t-1} = \emptyset$.

If we consider the case of full decay of past frustration (i.e., $\delta_i = 0$), we obtain the “slow play” version of the multistage setting in BD&S.⁴⁷

To extend the model for the multistage case ($L > 1$), define the flow of frustration experienced at the beginning of stage $\ell + 1 \in \{1, \dots, L\}$ of period t given history $h^{t-1, \ell}$ as⁴⁸

$$F_{i, \ell+1, t} \left[h^{t-1, \ell}, \alpha_i \right] := \left[\mathbb{E} \left[\Pi_{i, t}, \alpha_i | h^{t-1} \right] - \max_{a_i \in \mathcal{A}_i(h^{t-1, \ell})} \mathbb{E} \left[\Pi_{i, t}, \alpha_i | h^{t-1, \ell}, a_i \right] \right]^+,$$

while the stock of frustration is

$$F_i^{\ell+1, t} \left[h^{t-1, \ell}, \alpha_i \right] := F_{i, \ell+1, t} \left[h^{t-1, \ell}, \alpha_i \right] + \sum_{k=1}^{t-1} \delta_i^{t-k+1} F_{i, L, k} \left[h^k, \alpha_i \right]$$

where, for each $k \leq t - 1$, h^k denotes the k -period complete history that precedes $h^{t-1, \ell}$. This formula is coherent with our interpretation that, within period t , interaction takes place almost instantaneously and then players wait one unit of time before moving on to period $t + 1$. Hence, there is no within-period accumulation of frustration, just a within period accumulation of information that affects *via* updating players’ perception of goal obstruction. Frustration is only accumulated across periods.

Section summary Our definition of p-game consists of two main ingredients:

1. (trait- and) belief-dependent utilities of terminal histories $u_i : Z \times \Theta \times \Delta_1 \rightarrow \mathbb{R}$ ($i \in I$), which we interpret as experience utilities;
2. local (i.e., history-dependent) distortions $d_{i, h} : Z \times \Theta \times \Delta_1 \rightarrow \mathbb{R}$ ($i \in I, h \in H$), which help shape action tendencies.

According to our simplified assumptions, both depend on own and others’ *first-order* beliefs. Overall action tendencies are represented by local decision utilities $\bar{u}_{i, h} : \mathcal{A}_i(h) \times \Theta_i \times \Delta_{i, 2} \rightarrow \mathbb{R}$ ($i \in I, h \in H$) obtained as the summation of the conditional expectation of the experience utility and distortion for each action, given the player’s personal traits and second-order beliefs. The first ingredient, experience utility, may be derived from game-independent psy-utilities as per eq. (10), or defined directly on the game form, relying on its causal structure.

This allows us to provide a potentially useful classification of p-games according to the belief-dependent motivations they encode. We already discussed game-independence. Other interesting special cases obtain when distortions are absent ($d_{i, h} \equiv 0$) and all that matters is the *anticipation* of psychological states (which in turn may depend on anticipatory feelings), or the polar case when experience utility does not depend on beliefs and the latter only affect action tendencies *via* distortions (that is, $u_i : Z \times \Theta \rightarrow \mathbb{R}$). We call the former case “**mere anticipation**” and the latter “**mere action tendencies.**” With mere anticipation, preferences may satisfy own-plan

⁴⁷Note, however, that BD&S assume that payoffs are realized only at the end of the game.

⁴⁸Recal that $h^{t-1, \ell} = (h^{t-1}, h_i^\ell)$ is the concatenation of the $(t - 1)$ -period history h^{t-1} with the within-period history h_i^ℓ , and that $h^{t-1, 0} = h^{t-1}$.

independence (OPI) with the ensuing dynamic consistency properties. To illustrate, we present a non-exhaustive classification of some models.

	Game-independent	Game-dependent
Mere anticipation, no OPI	<i>Exp. – based ref. – dependence</i> <i>Anxiety and suspense</i>	
Mere anticipation, OPI	<i>Guilt aversion</i> <i>Intrinsic reputation</i>	<i>Regret avoidance</i>
Mere action tendencies	—	<i>Frustration & anger</i> <i>Reciprocity</i>

7 Solution concepts

In this section, we start with the formal definition of **rationality of player i** given his personal features $(\theta_i, \alpha_i) \in \Theta_i \times \Delta_{i,1}$ and his second-order belief $\beta_i \in \Delta_{i,2}$. This is a *conjunction of coherence conditions* concerning the triple (z, θ_i, β_i) . Specifically, the plan of i , uniquely derived from his belief β_i must give positive probability only to actions that maximize the “local” decision-utility functions $\bar{u}_{i,h}(\cdot; \theta_i, \beta_i)$ (**rational planning, RP**) and it must agree with the actual actions played by i on path z (**material consistency, MC**). Next, we introduce the concept of strong belief. Fix an event $F \subseteq Z \times \Theta_{-i} \times \Delta_{-i,1}$ (e.g., that i ’s co-players are rational). We say that i **strongly believes** F if his conditional second-order belief β_i over $Z \times \Theta_{-i} \times \Delta_{-i,1}$ gives probability 1 to F for each personal history $h_i \in H_i$ that does not contradict F . With this, we provide an algorithm, **strong-rationalizability**, characterizing the relevant implications of the epistemic hypotheses of rationality and common strong belief in rationality.⁴⁹

We also consider (pure, psychological) **self-confirming equilibrium (SCE)**, that is, a state at which every player best replies to his beliefs and these are confirmed by what he observes. Note, in a SCE, players might well hold wrong beliefs concerning co-players’ personal features or reactions to unchosen actions. SCE characterizes steady states of learning dynamics in games played recurrently, but—unlike rationalizability—it does not capture any form of strategic reasoning. By adding the confirmation property to strong rationalizability, we obtain **rationalizable SCE**.⁵⁰ Finally, we give a novel definition of **Bayesian sequential equilibrium (BSE)** for p-games, which generalizes and adapts the equilibrium concepts defined in B&D, Attanasi et al. (2016) and BD&S.

7.1 Rationality

In what follows, we maintain the assumption of coherence of first- and second-order beliefs, that is, $\alpha_i = \text{marg}_i(\beta_i)$, an obvious rationality condition (cf. Section 5.2). We now define all the other aspects of players’ rationality.

⁴⁹For applications of rationalizability to psychological games see Attanasi et al. (2013) and Battigalli et al. (2013).

⁵⁰The term “self-confirming equilibrium” was coined in the seminal article of Fudenberg & Levine (1993). See the survey by Battigalli et al. (1992) and the discussion in Battigalli et al. (2015) for references to the (rationalizable) SCE and related concepts.

Definition 3 *Player $i \in I$ plans rationally given (θ_i, β_i) if, for every $h \in H$,*

$$\text{supp } \hat{\sigma}_i(\alpha_i)(\cdot|h) \subseteq r_{i,h}(\theta_i, \beta_i), \quad (12)$$

where $\alpha_i = \text{marg}_i(\beta_i)$.

Rational planning (RP) is a coherence condition relating i 's plan to θ_i and his beliefs about others. We can interpret it as an *intra-personal equilibrium* (incentive compatibility) condition saying that planned actions maximize decision utility given “rational expectations” about future own behavior. RP also applies to players whose preferences are dynamically inconsistent, as long as players are sophisticated.⁵¹ The following **RP correspondence** of i associates each θ_i with the beliefs of i satisfying RP given θ_i :

$$\begin{aligned} RP_i : \quad \Theta_i &\rightrightarrows \Delta_{i,2}, \\ \theta_i &\mapsto \bigcap_{h \in H} \{\beta_i \in \Delta_{i,2} : \hat{\sigma}_i(\alpha_i)(r_{i,h}(\theta_i, \beta_i)|h) = 1\}. \end{aligned}$$

Under own-plan independence (OPI, Definition 2), the standard properties of the subjective expected utility model hold and rational plans are obtained by dynamic programming methods. More formally, expected experience utility can be expressed as a function of i 's plan and beliefs about co-players, denoted by $\beta_{i,-i}$:

$$E[u_i(\cdot, \theta_i, \text{marg}_i(\beta_i)), \beta_i|h] \equiv U_{\theta_i,h}(\hat{\sigma}_i(\text{marg}_i(\beta_i)), \beta_{i,-i}),$$

where $U_{\theta_i,h}$ is a function of i 's behavior strategy (starting at h) and i 's beliefs about others.⁵²

Remark 1 *Under OPI, player i plans optimally given (θ_i, β_i) if and only if*

$$E[u_i(\cdot, \theta_i, \text{marg}_i(\beta_i)), \beta_i|h] = \max_{\sigma_i \in \Sigma_i} U_{\theta_i,h}(\sigma_i, \beta_{i,-i})$$

for every $h \in H$. Furthermore, for every $h \in H$, the maximum is also attained by every pure continuation strategy in the “support” of $\hat{\sigma}_i(\text{marg}_i(\beta_i))$ [that is, $\prod_{h' \succeq h} \text{supp } \hat{\sigma}_i(\text{marg}_i(\beta_i))(\cdot|h')$].

In words, under OPI optimal planning is equivalent to sequential rationality. If instead OPI does not hold, such equivalence fails. Furthermore, for given beliefs about others, it may be impossible to find deterministic plans satisfying RP, as the following example due to BD&S shows (see also Section 6 in B&D).

Rational planning in the Ultimatum Minigame with Frustration and Anger Consider the one-period Ultimatum Minigame in Fig. 9 and assume that Bob is affected by frustration and anger. As shown in Fig. 10, Bob's decision utility at history (Gr) is

$$u_B(a_B; \theta_B, \beta_B) = \begin{cases} 1 - 3\theta_B [2\alpha_B(Fa|\emptyset) + \alpha_B(Ac|Gr)(1 - \alpha_B(Fa|\emptyset)) - 1]^+ & \text{if } a_B = Ac, \\ 0 & \text{if } a_B = Rj, \end{cases}$$

⁵¹See the example on anger in the Ultimatum Minigame.

⁵²Such expression relies on the chain rule and OAI. We omit the details.

where $\alpha_B = \text{marg}_B(\beta_B)$ and $\alpha_B(Ac|Gr) := \hat{\sigma}_B(\alpha_B)(Ac|Gr)$. Assume that $\theta_B = 1$ and that Bob initially expects a fair offer with probability $1/2$. One can check that Bob has no deterministic plan satisfying RP. In particular, if he initially plans to *Accept*, after the *Greedy* offer he wants to *Reject* because his initially expected payoff—hence his later frustration—is high enough to induce an angry reaction. If instead he plans to *Reject*, after the offer he wants to *Accept*, because his initially expected payoff—hence his later frustration—is too low. The only plan that satisfies the rationality condition is to accept with $2/3$ probability ($\alpha_B(Ac|Gr) = 2/3$).

Next we move to consistency between plan and actual behavior.

Definition 4 *Player $i \in I$ is **materially consistent (MC)** at (z, θ_i, α_i) if, for all $h \in H$ and $a_i \in \mathcal{A}_i(h)$,*

$$(h, a_i) \preceq z \implies \hat{\sigma}_i(\alpha_i)(a_i|h) > 0.$$

MC of player i imposes a restriction on the pair (z, β_i) : for every personal history $(h, a_i) \in H_i$ preceding terminal history z , the local plan at h derived from β_i must assign positive probability to the actually chosen action a_i . The **MC correspondence** of i associates each belief (second-order CPS) β_i with the set of paths z such that i satisfies MC given β_i :

$$\begin{aligned} Z_i^{MC} : \Delta_{i,2} &\rightrightarrows Z, \\ \beta_i &\mapsto \bigcap_{h \in H} \bigcap_{a_i \in \mathcal{A}_i(h)} \{z \in Z : (h, a_i) \preceq z \implies \hat{\sigma}_i(\text{marg}_i(\beta_i))(a_i|h) > 0\}. \end{aligned}$$

We can now define **rationality** of player i .

Definition 5 *Player $i \in I$ is **rational** at (z, θ_i, β_i) if he plans rationally given (θ_i, β_i) and is materially consistent at (z, θ_i, β_i) , that is, if*

$$(z, \theta_i, \beta_i) \in R_i := \{(z', \theta'_i, \beta'_i) \in Z \times \Theta_i \times \Delta_{i,2} : (z', \beta'_i) \in Z_i^{MC}(\beta'_i) \times RP_i(\theta'_i)\}.$$

Thus, player i is rational at (z, θ_i, β_i) if and only if (z, β_i) is a *fixed point* of the following intra-personal equilibrium correspondence

$$\begin{aligned} \Gamma_{\theta_i} : Z \times \Delta_{i,2} &\rightrightarrows Z \times \Delta_{i,2}, \\ (z, \beta_i) &\mapsto Z_i^{MC}(\beta_i) \times RP_i(\theta_i). \end{aligned}$$

Given this characterization of rationality, one can prove the existence of states in $Z \times \Theta_i \times \Delta_{i,2}$ at which i is rational by fixed-point methods if u_i is continuous.

7.2 Strong belief and restrictions of beliefs

The rationality concept defined in the previous section does not capture strategic thinking. We now proceed to model a form of strategic thinking based on the assumption that each player keeps assuming that his co-players are rational (and strategically sophisticated) as long as this is consistent with their observed behavior. This is a form of forward-induction reasoning (cf.

Battigalli & Siniscalchi 2002). Such strategic reasoning helps players infer—to some extent—what other players think from the rationalization of what they do. This is particularly important in p-games where players care about the beliefs of others as they matter intrinsically, not just instrumentally, to predict the likely consequences of own actions (cf. B&D). Forward induction is based on the notion of strong belief:

Definition 6 Fix a player $i \in I$ and a second-order belief $\beta_i \in \Delta_{i,2}$. We say that β_i **strongly believes** event $F \subseteq Z \times \Theta_{-i} \times \Delta_{-i,1}$ if, for every personal history $h_i \in H_i$,

$$F \cap [h_i] \neq \emptyset \implies \beta_i(F|h_i) = 1,$$

where $[h_i] := Z(h_i) \times \Theta_{-i} \times \Delta_{-i,1}$.

With this, we are able to express hypotheses about strategic thinking, such as the assumption that player i strongly believes in the rationality of his co-players.

In applications, it is often the case that beliefs of some (or all) players are required to satisfy restrictions that are not derived by strategic thinking. Such restrictions may depend parametrically on own personal traits. For example, we may want to assume that each player thinks the personal traits of others are similar to his own. Our framework allows us to express such restrictions. Formally, we consider a profile of *nonempty compact-valued* correspondences $\hat{\Delta} := \left(\hat{\Delta}_i : \Theta_i \rightrightarrows \Delta_{i,2} \right)_{i \in I}$ such that, for each player $i \in I$, $\hat{\Delta}_i$ associates each θ_i with a corresponding set of second-order beliefs $\hat{\Delta}_i(\theta_i) \subseteq \Delta_{i,2}$. We say that the belief of i **satisfies restrictions $\hat{\Delta}$** at (z, θ_i, β_i) if $\beta_i \in \hat{\Delta}_i(\theta_i)$. These restrictions may represent commonly understood probabilities of chance moves, or other assumptions suggested by the application at hand like, e.g., the symmetry assumptions in the analysis of rationalizable deception under guilt aversion in Battigalli et al. (2013).

7.3 Strong rationalizability

We can finally present the strong rationalizability algorithm. The procedure we propose entails an *iterated elimination of utility-relevant states*. Importantly, this *is not* a procedure on observables since, for example, the profile of first-order beliefs α_i of player i cannot be directly observed.⁵³ Nevertheless, in many experimental settings, it is possible to elicit the first-order beliefs of players in order to verify the accuracy of the predictions provided by strong rationalizability.

Definition 7 Fix restrictions $\hat{\Delta}$ and consider the following procedure

(Step 0) Let $P_{c, \hat{\Delta}}^0 := Z \times \Theta_c$; for every $i \in I$, let $P_{i, \hat{\Delta}}^0 := Z \times \Theta_i \times \Delta_{i,1}$ and $\mathbf{P}_{-i, \hat{\Delta}}^0 := Z \times \Theta_{-i} \times \Delta_{-i,1}$;

(Step $n > 0$) Let $P_{c, \hat{\Delta}}^n := P_{c, \hat{\Delta}}^{n-1} = Z \times \Theta_c$; for every $i \in I$, $P_{i, \hat{\Delta}}^n$ is the set of triples $(z, \theta_i, \alpha_i) \in Z \times \Theta_i \times \Delta_{i,1}$ such that there exists $\beta_i \in \hat{\Delta}_i(\theta_i)$ satisfying

⁵³Of course, it is always possible to consider the projection of the set of rationalizable states in $Z \times \Theta \times \Delta_1$ onto the space of observables.

1. **Coherence:** $\text{marg}_i(\beta_i) = \alpha_i$;
2. **Rationality:** Player i is rational at (z, θ_i, β_i) ;
3. **Strong belief:** For every $m \in \{1, \dots, n-1\}$, β_i strongly believes $\mathbf{P}_{-i, \hat{\Delta}}^m$.

Moreover, let

$$\mathbf{P}_{-i, \hat{\Delta}}^n := \bigcap_{j \in I_c \setminus \{i\}} \left(P_{j, \hat{\Delta}}^n \times \Theta_{-j} \times \Delta_{-j, 1} \right)$$

and

$$\mathbf{P}_{\hat{\Delta}}^n := \bigcap_{i \in I_c} \left(P_{i, \hat{\Delta}}^n \times \Theta_{-i} \times \Delta_{-i, 1} \right).$$

We say that player i is n -th order rational at (z, θ, α) if this triple survives the n -th iteration of the above procedure, that is, if $(z, \theta, \alpha) \in \mathbf{P}_{\hat{\Delta}}^n$. It can be checked by inspection of the recursive definition that $\mathbf{P}_{\hat{\Delta}}^n \subseteq \mathbf{P}_{\hat{\Delta}}^{n-1}$ for every $n \in \mathbb{N}$. Finally, we define $\mathbf{P}_{\hat{\Delta}}^\infty := \bigcap_{n \in \mathbb{N}} \mathbf{P}_{\hat{\Delta}}^n$ and say that

(z, θ, α) is **strongly $\hat{\Delta}$ -rationalizable** if $(z, \theta, \alpha) \in \mathbf{P}_{\hat{\Delta}}^\infty$.

BC&S provide an epistemic justification for the algorithm of Definition 7. In particular, they show that, for every $n \in \mathbb{N}$, $\mathbf{P}_{\hat{\Delta}}^n$ characterizes the utility-relevant implications of the epistemic hypotheses of rationality and n -mutual strong belief in rationality given the restrictions $\hat{\Delta}$. Under own-plan independence of psy-utility (OPI), $\mathbf{P}_{\hat{\Delta}}^\infty$ characterizes the utility-relevant implications of rationality and common strong belief in rationality given the restrictions $\hat{\Delta}$ (for an illustration, see the analysis of the Trust Minigame in Section 7.5). Furthermore, nonemptiness can be proved for a class of restrictions, given OPI and continuous psy-utilities.⁵⁴

Result 1 *Suppose that the restrictions $\hat{\Delta}$ only concern beliefs about the co-players' personal traits, OPI holds, and players' experience utilities $(u_i, i \in I)$ are continuous in beliefs. Then, for every $n \in \mathbb{N} \cup \{\infty\}$, $\mathbf{P}_{\hat{\Delta}}^n$ is nonempty.*

7.4 Self-confirming equilibrium (SCE)

To ease the exposition, in the present section we assume that there are *no chance moves* and that psy-utility is *own-plan independent*. Consider now the following simplified situation. The same set of agents play the same one-period (possibly multi-stage) game among themselves for a very large number of periods. Furthermore, these agents are very impatient; therefore, they do not care about the impact of their current period behavior on the behavior and beliefs of their co-players in future periods, that is, they just maximize their one-period expected utility. We want to study stable behavior and beliefs. Stability obtains when updated beliefs and best-reply behavior converge to a limit and do not change any more. Limit behavior is described by a path z of the one-period game. Assuming that players observe ex post the realized path, it must be

⁵⁴On p-games, see also BC&S and Section 5 in B&D. There are also other notions of rationalizability for dynamic games with different epistemic justifications. For example, Perea (2014) shows that common belief in future rationality yields a notion of “backward rationalizability” (see also Battigalli & De Vito 2018), and Ben-Porath (1997) shows that common *initial* belief in rationality justifies a much weaker notion of rationalizability (see also Battigalli & Siniscalchi 1999).

the case that in the limit they assign probability 1 to the actions played on path z . Thus, beliefs about the behavior of others must be confirmed in the following sense.

Definition 8 *Player $i \in I$ has **confirmed beliefs** about others at (z, θ_i, β_i) if, for all $h \in H$ and $\mathbf{a}_{-i} \in \mathcal{A}_{-i}(h)$,*

$$(h, \mathbf{a}_{-i}) \preceq z \implies \hat{\sigma}_{i,-i}(\alpha_i)(\mathbf{a}_{-i}|h) = 1,$$

where $\alpha_i = \text{marg}_i(\beta_i)$.

Since we assumed that psy-utility satisfies OPI, we may also assume without essential loss of generality that players have deterministic plans (see Section 7.1). Thus, we consider the following strengthening of the MC condition:

Definition 9 *Player $i \in I$ is **deterministically materially consistent (DMC)** at (z, θ_i, β_i) if, for all $h \in H$ and $a_i \in \mathcal{A}_i(h)$,*

$$(h, a_i) \preceq z \implies \hat{\sigma}_i(\alpha_i)(a_i|h) = 1,$$

where $\alpha_i = \text{marg}_i(\beta_i)$.

We can compactly express DMC and confirmed beliefs about others as follows:

Definition 10 *Player $i \in I$ has **path-consistent beliefs** at (z, θ_i, β_i) if, for all $h \in H$ and $\mathbf{a} \in \mathcal{A}(h)$,*

$$(h, \mathbf{a}) \preceq z \implies \hat{\sigma}_{i,I}(\alpha_i)(\mathbf{a}|h) = 1,$$

where $\alpha_i = \text{marg}_i(\beta_i)$.

We can now define (psychological) SCE (cf. Fudenberg & Levine 1993).

Definition 11 *Fix belief restrictions $\hat{\Delta}$. Triple $(z, \boldsymbol{\theta}, \boldsymbol{\alpha})$ is an **SCE** given $\hat{\Delta}$ if, for every $i \in I$, there exists $\beta_i \in \hat{\Delta}_i(\theta_i)$ such that*

1. **Coherence:** $\text{marg}_i(\beta_i) = \alpha_i$;
2. **Rationality and path consistency:** *Player i is rational and path consistent at (z, θ_i, β_i) .*

It is clear that, whenever we let $\hat{\Delta}_i(\theta_i) = \Delta_{i,2}$ for all $i \in I$ and $\theta_i \in \Theta_i$, then we get a “restriction-free” definition of SCE.

Note that the foregoing equilibrium concept does *not* capture endogenous restrictions of beliefs derived from strategic reasoning. Indeed, we can make sense of this definition even if we do not assume that players’ parameterized utility functions are commonly known. Suppose now that the parameterized utility functions are commonly known (i.e., it is common knowledge how $\boldsymbol{\theta}$ maps to belief-dependent preferences). Then, it makes sense to define a refinement of self-confirming equilibrium that takes into account strategic reasoning based on such common knowledge (cf. Rubinstein & Wolinsky 1994 and Battigalli & Guaitoli 1997).

Definition 12 Fix restrictions $\hat{\Delta}$ and consider the following procedure

(Step 0) For every $i \in I$, let $Q_{i,\hat{\Delta}}^0 := Z \times \Theta_i \times \Delta_{i,1}$ and $\mathbf{Q}_{-i,\hat{\Delta}}^0 := Z \times \Theta_{-i} \times \Delta_{-i,1}$;

(Step $n > 0$) For every $i \in I$, $Q_{i,\hat{\Delta}}^n \subseteq Z \times \Theta_i \times \Delta_{i,1}$ is the set of triples (z, θ_i, α_i) such that there exists $\beta_i \in \hat{\Delta}_i(\theta_i)$ satisfying

1. **Coherence:** $\text{marg}_i(\beta_i) = \alpha_i$;
2. **Rationality and path consistency:** Player i is rational and path consistent at (z, θ_i, β_i) ;
3. **Strong belief:** For every $m \in \{1, \dots, n-1\}$, β_i strongly believes $\mathbf{Q}_{-i,\hat{\Delta}}^m$.

Moreover, let

$$\mathbf{Q}_{-i,\hat{\Delta}}^n := \bigcap_{j \in I \setminus \{i\}} \left(Q_{j,\hat{\Delta}}^n \times \Theta_{-j} \times \Delta_{-j,1} \right),$$

$$\mathbf{Q}_{\hat{\Delta}}^n := \bigcap_{i \in I} \left(Q_{i,\hat{\Delta}}^n \times \Theta_{-i} \times \Delta_{-i,1} \right),$$

and $\mathbf{Q}_{\hat{\Delta}}^\infty := \bigcap_{n \in \mathbb{N}} \mathbf{Q}_{\hat{\Delta}}^n$. We say that $(z, \theta, \alpha) \in Z \times \Theta \times \Delta_1$ is a **strongly $\hat{\Delta}$ -rationalizable SCE** if and only if $(z, \theta, \alpha) \in \mathbf{Q}_{\hat{\Delta}}^\infty$.

As in the case of strong rationalizability, we can provide an epistemic foundation for the algorithm of Definition 12.⁵⁵ Since we are considering a “pure” version of SCE without actual randomization, we cannot give general existence conditions.

7.5 Examples about solution concepts with foundations

We analyze two psychological games—the Trust Minigame with simple Guilt and the Ultimatum Minigame with Simple Anger—through the solution concepts presented above. Let us emphasize that these are just numerical examples illustrating the solution concepts. In particular, the behavioral implications crucially depend on the assumed restrictions on beliefs about traits.

Trust Minigame with simple guilt Consider the psychological game in Fig. 5 and let $\Theta_B = \{\underline{\theta}, \bar{\theta}\}$, with $\underline{\theta} < 1$, $\bar{\theta} > 2$. Since Ann is commonly known to be selfish, $\Theta_A = \{0\}$, and we omit θ_A from the notation. We assume no belief restriction for Bob, whereas Ann is assumed to believe that the high-guilt type of Bob is more likely than the low-guilt type:

$$\hat{\Delta}_A = \left\{ \beta_A \in \Delta_{A,2} : \text{marg}_{\Theta_B} \beta_A (\{\bar{\theta}\} | \emptyset) \geq \frac{1}{2} + \varepsilon \right\}, \quad (13)$$

where $0 < \varepsilon < 1/2$ is a fixed parameter. Recall that $Z = \{(Out), (In, Sh), (In, Tk)\}$. We use the same abbreviations of Section 2, such as $\alpha_A^{Sh} = \alpha_A(\{Sh\} | In) = \alpha_A(\{Sh\} \times \Theta_B | In)$, and D_A for Ann’s (belief-dependent) disappointment at (In, Tk) . We analyze this game with strong rationalizability and SCE.

⁵⁵The formal result and proof (which relies on own-plan independence) are available upon request.

Strong rationalizability The first three iterations of strong $\hat{\Delta}$ -rationalizability give the key predictions:

1. The first iteration of the algorithm in Definition 7 deletes the utility-relevant states that violate rationality, or the belief restrictions. Focusing on Ann, since she maximizes her expected material payoff, $(z, \alpha_A) \in P_{A, \hat{\Delta}}^1$ iff (if and only if) restriction (13) holds, and either $\alpha_A^{Sh} < 1/2$ with $z = (Out)$ and $\alpha_A^{In} = 0$, or $\alpha_A^{Sh} > 1/2$ with $z \in \{(In, Sh), (In, Tk)\}$ and $\alpha_A^{In} = 1$, or $\alpha_A^{Sh} = 1/2$ (and material consistency holds). In particular, In is consistent with Ann's rationality and belief restrictions, and $((In, a_B), \alpha_A) \in P_{A, \hat{\Delta}}^1$ for some $a_B \in \{Sh, Tk\}$ only if $E[\pi_A, \alpha_A | \emptyset] \geq 1$. Also, the low-guilt type of Bob wants to *Take*: $((In, a_B), \theta, \alpha_B) \in P_{A, \hat{\Delta}}^1$ only if $a_B = Tk$, $\alpha_B^{Sh} = 0$.
2. Since Bob strongly believes $P_{A, \hat{\Delta}}^1$, which is consistent with In , $(z, \theta_B, \alpha_B) \in P_{B, \hat{\Delta}}^2$ only if there is some β_B inducing α_B such that $\beta_B(E[\pi_A, \tilde{\alpha}_A | \emptyset] \geq 1 | In) = 1$. This implies $E[D_A, \beta_B | In] \geq 1$. Guilt aversion (Eq. 1) implies that Bob wants to *Share* (resp. *Take*) if $\theta_B = \bar{\theta}$ (resp. $\theta_B = \underline{\theta}$).
3. Since Ann strongly believes $P_{B, \hat{\Delta}}^2$, $\alpha_A^{Sh} = \text{marg}_{\Theta_B} \beta_A(\bar{\theta} | \emptyset)$, that is, she believes that Bob would *Share* iff $\theta_B = \bar{\theta}$. Thus, belief restriction (13) implies that $\alpha_A^{Sh} > 1/2$ and Ann wants to go *In*: $(z, \alpha_A) \in P_{A, \hat{\Delta}}^3$ only if $z = (In, a_B)$ ($a_B \in \{Sh, Tk\}$) and $\alpha_A^{In} = 1$.

SCE There are two cases:

- If Bob has low guilt aversion ($\theta_B = \underline{\theta}$), only path (*Out*) is consistent with SCE, because a low-guilt Bob would *Take* if given the opportunity. In particular, all triples $(Out, \underline{\theta}, \alpha)$ with $\alpha_A \in \hat{\Delta}_A$, $\alpha_A^{Sh} \leq 1/2$, $\alpha_B^{Sh} = 0$, and $\alpha_A^{In} = \alpha_B^{In} = 0$ are part of an SCE, where the last two equalities follow from path consistency.
- If Bob is highly guilt averse ($\theta_B = \bar{\theta}$), then both paths (*Out*) and (In, Sh) are consistent with SCE. Bob's best reply depends on his conditional belief $\beta_B(\cdot | In)$, which is *not* pinned down by the SCE conditions. Thus, all triples $((In, Sh), \bar{\theta}, \alpha)$ with $\alpha_A \in \hat{\Delta}_A$, $\alpha_A^{In} = \alpha_B^{In} = 1$, and $\alpha_A^{Sh} = \alpha_B^{Sh} = 1$, and $\alpha_A^{In} = \alpha_B^{In} = 0$ (by path consistency) are SCE's. Path (In, Tk) instead is inconsistent with SCE: for every triple $((In, Tk), \theta_B, \alpha)$, if path consistency holds, then $\alpha_A^{In} = 1$, $\alpha_A^{Sh} = 0$ and Ann cannot be rational.

With the given belief restrictions (Ann believes that $\bar{\theta}$ is more likely than $\underline{\theta}$), there is no strongly $\hat{\Delta}$ -rationalizable SCE if Bob has low guilt aversion. Indeed, a forward-induction argument similar to the one given for $\hat{\Delta}$ -rationalizability implies that Ann goes *In*, but then Bob *Takes* and Ann's beliefs cannot be confirmed (path consistency is violated). Only $((In, Sh), \bar{\theta})$ is consistent with strongly $\hat{\Delta}$ -rationalizable SCE.

Ultimatum Minigame with Simple Anger Consider the psychological game in Fig. 5 and let $\Theta_B = \{0, 1\}$. As noted by BD&S, since Ann is a first mover she cannot feel frustrated and her anger trait is irrelevant (see Section 6.2). Hence, we do not consider θ_A . We assume no belief

restriction for Bob, whereas Ann is assumed to believe that the probability of high-anger type of Bob is less than $1/3$:

$$\hat{\Delta}_A = \left\{ \beta_A \in \Delta_{A,2} : \text{marg}_{\Theta_B} \beta_A (\{1\} | \emptyset) \leq \frac{1}{3} - \varepsilon \right\},$$

where $0 < \varepsilon < 1/3$ is a fixed parameter. Recall that $Z = \{(Fa), (Gr, Ac), (Gr, Rj)\}$. We use the same abbreviations of Section 2, such as $\alpha_A^{Ac} = \alpha_A(Ac|Gr)$.

Strong rationalizability

1. Focusing on Ann, since she maximizes her expected material payoff, $(z, \alpha_A) \in P_{A, \hat{\Delta}}^1$ iff either $\alpha_A^{Ac} < 2/3$ with $z = (Fa)$ and $\alpha_A^{Gr} = 0$, or $\alpha_A^{Ac} > 2/3$ with $z \in \{(Gr, Ac), (Gr, Rj)\}$ and $\alpha_A^{Gr} = 1$, or $\alpha_A^{Ac} = 2/3$ (and material consistency holds). In particular, Gr is consistent with Ann's rationality and $((Gr, a_B), \alpha_A) \in P_{A, \hat{\Delta}}^1$ for some $a_B \in \{Ac, Rj\}$ only if $E[\pi_A, \alpha_A | \emptyset] \geq 1$. Conditional on Gr , when unaffected by anger (i.e., $\theta_B = 0$), Bob wants to *Accept*. When prone to anger ($\theta_B = 1$) Bob's rational plan depends on his belief α_B^{Fa} :

$$\hat{\alpha}_B^{Ac}(\alpha_B^{Fa}) = \begin{cases} 1 & \text{if } \alpha_B^{Fa} \leq \frac{1}{3} \\ \frac{\frac{4}{3} - 2\alpha_B^{Fa}}{1 - \alpha_B^{Fa}} & \text{if } \alpha_B^{Fa} \in (\frac{1}{3}, \frac{2}{3}) \\ 0 & \text{if } \alpha_B^{Fa} \geq \frac{2}{3} \end{cases}$$

where, for every $\alpha_B^{Fa} \in [0, 1]$, $\hat{\alpha}_B^{Ac}(\alpha_B^{Fa})$ denotes the probability to *Accept* prescribed by Bob's optimal plan given α_B^{Fa} . Therefore, $((Gr, a_B), 1, \alpha_B) \in P_{B, \hat{\Delta}}^1$ iff $\alpha_B^{Fa} < 2/3$ with $a_B = Ac$ and $\alpha_B^{Ac} = \min\{(\frac{4}{3} - 2\alpha_B^{Fa}) / (1 - \alpha_B^{Fa}), 1\}$, or $\alpha_B^{Fa} > 1/3$ with $a_B = Rj$ and $\alpha_B^{Ac} = \max\{(\frac{4}{3} - 2\alpha_B^{Fa}) / (1 - \alpha_B^{Fa}), 0\}$, or $\alpha_B^{Fa} = 2/3$ (and material consistency holds).

2. Since $P_{A, \hat{\Delta}}^1$ allows for both offers by Ann, Bob's strong belief in $P_{A, \hat{\Delta}}^1$ does not yield new relevant implications for Bob's beliefs and behavior. Instead, Ann's belief in Bob's rationality, and the assumption (belief restrictions) that she assigns low probability to the high-anger type imply that she makes the *Greedy* offer: $(z, \alpha_A) \in P_{A, \hat{\Delta}}^2$ only if $z = (Gr, a_B)$ ($a_B \in \{Ac, Rj\}$), $\alpha_A^{Ac} \geq 2/3 + \varepsilon$, and $\alpha_A^{Gr} = 1$.
3. Bob's belief in $P_{A, \hat{\Delta}}^2$ implies that he expects the greedy offer and is not frustrated by it. Thus, he *Accepts* even if prone to anger: $((Gr, a_B), 1, \alpha_B) \in P_{B, \hat{\Delta}}^3$ only if $a_B = Ac$, $\alpha_B^{Fa} = 0$, $\alpha_B^{Ac} = 1$.

SCE

- Path (Fa) is always consistent with SCE. Indeed, whenever $\alpha_A^{Ac} \leq 2/3$, rationality implies that Ann makes the *Fair* offer. With this, Ann and Bob's conjectures will never be falsified.
- Path (Gr, Ac) is always consistent with SCE. Indeed, in this case path consistency implies

$$\alpha_A^{Gr} = \alpha_B^{Gr} = \alpha_B^{Ac} = \alpha_A^{Ac} = 1.$$

Therefore, both Ann and Bob also satisfy rational planning. Indeed, expecting the Greedy offer with probability 1, Bob is never frustrated, regardless of his personal trait. Path (Gr, Rj) is inconsistent with SCE because path consistency contradicts rationality. Note that behavioral implications of SCE are not affected by anger traits and belief restrictions.

In this example, strongly $\hat{\Delta}$ -rationalizable SCE works pretty much like strong $\hat{\Delta}$ -rationalizability.

The following table summarizes the main predictions of the solution concepts in the two p-games, which of course depend on the assumed restrictions on beliefs.

	str. $\hat{\Delta}$ -rationalizability	$\hat{\Delta}$ -SCE	str. $\hat{\Delta}$ -rationalizable SCE
<i>TmG with Guilt</i> $\hat{\Delta} : \alpha_A(\bar{\theta}) \geq 1/2 + \varepsilon$	(In, Sh) if $\theta_B = \bar{\theta}$ (In, Tk) if $\theta_B = \underline{\theta}$	(Out) or (In, Sh) if $\theta_B = \bar{\theta}$ (Out) if $\theta_B = \underline{\theta}$	(In, Sh) if $\theta_B = \bar{\theta}$ \emptyset (\nexists $\hat{\Delta}$ -rat. SCE) if $\theta_B = \underline{\theta}$
<i>UmG with Anger</i> $\hat{\Delta} : \alpha_A(\theta_B = 1) \leq 1/3 - \varepsilon$	(Gr, Ac)	(Fa) or (Gr, Ac)	(Gr, Ac)

7.6 Bayesian sequential equilibrium (BSE)

Rationalizability and (rationalizable) SCE are non-standard solution concepts with a rigorous epistemic or learning foundation. Now we move on to an extension for p-games of a solution concept with no deep foundation (to the best of our knowledge), but with tradition on its side: K&W's sequential equilibrium (SE), a slight coarsening of Selten's (1975) trembling-hand perfect equilibrium concept. We will argue that earlier extensions to p-games of traditional equilibrium concepts are special cases of the solution concept proposed here, called Bayesian SE, or BSE. The advantage of this perspective in PGT is that it yields the hierarchies of beliefs that PGT takes as arguments of the utility functions.

To simplify the exposition, here we assume that Θ_c is a singleton, i.e., there is *distributed knowledge* of θ . Thus, we can neglect the chance player c and write $\Theta = \prod_{i \in I} \Theta_i$. We further simplify by neglecting the role of time: we assume that we have a *one-period multistage game*. Both assumptions can be removed by increasing notational complexity.

Harsanyi's Bayesian equilibrium and belief hierarchies Our analysis is based on Harsanyi's (1967-68) seminal work on games with incomplete information, which we have to review. Let $\Delta^\circ(X)$ denote the set of full-support probability measures over a *finite* set X . Harsanyi defined and analyzed equilibria of games with incomplete information starting with the following ingredients: a game form and a profile of parametrized (non-psychological) utility functions $(u_i : Z \times \Theta \rightarrow \mathbb{R})_{i \in I}$. Recognizing that beliefs about Θ and beliefs about beliefs are both essential for equilibrium analysis, he proposed to add to such ingredients a Θ -based **type structure**

$$\langle I, (\Theta_i, \mathcal{E}_i, \eta_i)_{i \in I} \rangle$$

that provides an implicit representation of hierarchical beliefs about Θ . Specifically, for player $i \in I$, \mathcal{E}_i is a finite (nonempty) set of **epistemic types**, and $\mathcal{T}_i := \Theta_i \times \mathcal{E}_i$ denotes i 's set of **types** à la Harsanyi. The (finite) set of profiles of types is $\mathcal{T} := \prod_{i \in I} \Theta_i \times \mathcal{E}_i$, and each player $i \in I$ is

endowed with a *strictly positive* exogenous **prior** belief $\eta_i \in \Delta^\circ(\mathcal{T})$.⁵⁶ Similarly, we denote with $\mathcal{T}_{-i} := \prod_{j \in I \setminus \{i\}} \Theta_j \times \mathcal{E}_j$ the (finite) set of profiles of types of player i 's co-players.

The prior η_i may be interpreted as the belief about exogenous unknowns held by i in a hypothetical ex ante stage in which he does not yet know his type,⁵⁷ but there is no need to give an independent meaning to η_i , it is just a conveniently compact way to encode—*via* conditioning—the exogenous belief of each type $\tau_i = (\theta_i, \varepsilon_i)$ about the types of other players. Thus, for each $i \in I$, we derive a belief map

$$\begin{aligned} \varphi_i : \mathcal{T}_i &\rightarrow \Delta^\circ(\mathcal{T}_{-i}), \\ \tau_i &\mapsto \frac{\eta_i(\tau_i, \cdot)}{\eta_i(\{\tau_i\} \times \mathcal{T}_{-i})}. \end{aligned}$$

It follows that, each type τ_i is associated with a coherent hierarchy of exogenous beliefs. To explain it rigorously, we first introduce preliminary notation. Fix a function $\varphi : \mathcal{X} \rightarrow \mathcal{Y}$, where \mathcal{X} and \mathcal{Y} are, so far, abstract sets. Then, every finite-support probability measure μ on \mathcal{X} induces a corresponding finite-support probability measure on \mathcal{Y} , denoted by $\mu \circ \varphi^{-1}$, according to the following **pushforward** formula:⁵⁸

$$(\mu \circ \varphi^{-1})(E) := \mu(\varphi^{-1}(E)) = \sum_{y \in E \cap \varphi(\text{supp } \mu)} \mu(\varphi^{-1}(y)).$$

The hierarchy of exogenous beliefs of type τ_i is a sequence

$$\varphi_i^\infty(\tau_i) = (\varphi_{i,n}(\tau_i))_{n=1}^\infty,$$

where the exogenous first-order belief is

$$\varphi_{i,1}(\tau_i) = \text{marg}_{\Theta_{-i}} \varphi_i(\tau_i) \in \Delta(\Theta_{-i}).$$

With this, for each $i \in I$, we can define the auxiliary map

$$\begin{aligned} \bar{\varphi}_{-i,1} : \mathcal{T}_{-i} &\rightarrow \prod_{j \neq i} \Theta_j \times \Delta(\Theta_{-j}), \\ (\theta_{-i}, \varepsilon_{-i}) &\mapsto (\theta_j, \varphi_{j,1}(\theta_j, \varepsilon_j))_{j \neq i} \end{aligned}$$

associating each profile of co-players' types with the corresponding profile of personal traits and first-order beliefs. Then, the exogenous second-order belief of each type τ_i is

$$\varphi_{i,2}(\tau_i) = \varphi(\tau_i) \circ \bar{\varphi}_{-i,1}^{-1} \in \Delta\left(\prod_{j \neq i} \Theta_j \times \Delta(\Theta_{-j})\right).$$

The second-order belief hierarchy $\varphi_i^2(\tau_i) = (\varphi_{i,1}(\tau_i), \varphi_{i,2}(\tau_i))$ is necessarily coherent, i.e.,

$$\text{marg}_{\Theta_{-i}} \varphi_{i,2}(\tau_i) = \varphi_{i,1}(\tau_i)$$

⁵⁶The assumption that the prior is strictly positive ($\eta_i(\tau) > 0$ for each $\tau \in \mathcal{T}$) is made for simplicity.

⁵⁷We call **exogenous** any belief about exogenous unknowns, either unconditional, or conditional on exogenous events. Otherwise, a belief is called **endogenous**.

⁵⁸The definition for arbitrary probability measures is similar, but it involves measure-theoretic technicalities.

by construction. We can similarly derive all the finite hierarchies of exogenous beliefs

$$\varphi_i^n(\tau_i) = (\varphi_{i,1}(\tau_i), \dots, \varphi_{i,n}(\tau_i))$$

as well as the infinite hierarchy $\varphi_i^\infty(\tau_i)$, and all these hierarchies are coherent by construction. Since we use only beliefs up to the second order, we need not spell out the details.

Three observations are worth noting.

1. Harsanyi's approach gives no guidance, or discipline, about the specification of the sets \mathcal{E}_i of epistemic types and the belief maps φ_i . For example, each set \mathcal{E}_i could be a singleton, thus making beliefs depend (only) on personal traits. A kind of polar case obtains when Θ is singleton, i.e., there is complete information, and yet multiple epistemic types, i.e., \mathcal{E}_i has at least two elements for some $i \in I$. This makes Harsanyi's approach at the same time baffling and very flexible. We take advantage of its flexibility.
2. According to Harsanyi's Bayesian equilibrium, a player's behavior depends on his type according to a behavior map (decision rule) ς_i , and each type of each player behaves so as to maximize expected utility given his exogenous belief under the (traditional) equilibrium assumption that he has a correct conjecture about the behavior maps ς_{-i} . In particular, a (pure) Bayesian equilibrium of a one-stage standard game is a profile $(\varsigma_i : \mathcal{T}_i \rightarrow A_i)_{i \in I}$ of (pure) **behavior maps** such that, for all $i \in I$ and $(\theta_i, \varepsilon_i) \in \mathcal{T}_i$,

$$\varsigma_i(\theta_i, \varepsilon_i) \in \arg \max_{a_i \in A_i} \mathbb{E} \left[u_i \left(a_i, \varsigma_{-i} \left(\tilde{\theta}_{-i}, \tilde{\varepsilon}_{-i} \right), \theta_i, \tilde{\theta}_{-i} \right), \varphi_i(\theta_i, \varepsilon_i) \right]$$

(the definition can be generalized to allow for randomized behavior maps, as shown below for p-games). The key observation here is that the profile of behavior maps $\varsigma = (\varsigma_i)_{i \in I}$ yields a **hierarchy of endogenous beliefs**⁵⁹

$$\beta_i^\infty(\tau_i, \varsigma) = (\beta_{i,n}(\tau_i))_{n=1}^\infty$$

for each type $\tau_i \in \mathcal{T}_i$ of each $i \in I$. To see this, for each i , define the auxiliary map

$$\begin{aligned} \bar{\varsigma}_{-i} : \quad \mathcal{T}_{-i} &\quad \rightarrow \quad \mathbf{A}_{-i} \times \Theta_{-i}, \\ (\theta_{-i}, \varepsilon_{-i}) &\quad \mapsto \quad (\varsigma_j(\theta_j, \varepsilon_j), \theta_j)_{j \neq i}. \end{aligned}$$

Then, the first-order belief of each τ_i about co-players' traits *and actions* is

$$\beta_{i,1}(\tau_i, \varsigma) = \varphi_i(\tau_i) \circ \bar{\varsigma}_{-i}^{-1} \in \Delta_{i,1} := \Delta(\mathbf{A}_{-i} \times \Theta_{-i}).$$

With this, we obtain, for each i , another auxiliary map:

$$\begin{aligned} \bar{\varsigma}_{-i,1} : \quad \mathcal{T}_{-i} &\quad \rightarrow \quad \mathbf{A}_{-i} \times \Theta_{-i} \times \Delta_{-i,1}, \\ (\theta_{-i}, \varepsilon_{-i}) &\quad \mapsto \quad (\varsigma_j(\theta_j, \varepsilon_j), \theta_j, \beta_{j,1}(\theta_j, \varepsilon_j, \varsigma))_{j \neq i} \end{aligned}$$

⁵⁹Since here we consider beliefs of many orders, we let $\beta_{i,n}$ denote the generic belief of order n . In particular, we write $\beta_{i,1}$ instead of α_i for first-order beliefs, and $\beta_{i,2}$ instead of β_i for second-order beliefs.

and derive the second-order belief of each type τ_i about the actions, traits, and first-order beliefs of the co-players as follows:

$$\beta_{i,2}(\tau_i, \varsigma) = \varphi_i(\tau_i) \circ \bar{\varsigma}_{-i,1}^{-1} \in \Delta_{i,2} := \Delta(\mathbf{A}_{-i} \times \Theta_{-i} \times \Delta_{-i,1}).$$

Higher-order endogenous beliefs are derived in a similar way. Again, we omit the details because we do not use them, and we henceforth revert to denoting first-order beliefs with α and second-order beliefs with β ; thus,

$$\alpha_i(\tau_i, \varsigma) := \beta_{i,1}(\tau_i, \varsigma) \text{ and } \beta_i(\tau_i, \varsigma) := \beta_{i,2}(\tau_i, \varsigma).$$

Like the exogenous hierarchies, also the hierarchies of endogenous beliefs are coherent; in particular,

$$\alpha_i(\tau_i, \varsigma) = \text{marg}_{\mathbf{A}_{-i} \times \Theta_{-i}} \beta_i(\tau_i, \varsigma).$$

3. With this construction and equilibrium concept, distinct types with the same exogenous hierarchical beliefs may behave differently in equilibrium. For example, if there is *complete information*, Θ is a singleton ($\Theta = \{\bar{\theta}\}$) and exogenous hierarchies of beliefs are trivial: the first-order belief of each type $(\bar{\theta}_i, \varepsilon_i) \in \{\bar{\theta}_i\} \times \mathcal{E}_i$ assigns probability 1 to $\bar{\theta}_{-i}$, the second-order belief assigns probability 1 to $\bar{\theta}_{-i}$ and the profile of such point beliefs of the co-players, and so on. Yet, there may be multiple epistemic types for some players, and a Bayesian equilibrium may be different from a Nash equilibrium of the complete-information game: indeed, it is a *correlated equilibrium* (Aumann, 1974) of the complete-information game.

Bayesian p-games and equilibria Suppose we have a *one-stage* p-game with utilities that depends only on initial beliefs:

$$\langle I, (A_i, \Theta_i, u_i : \mathbf{A} \times \Theta \times \Delta_1 \rightarrow \mathbb{R})_{i \in I} \rangle,$$

where $\Delta_{i,1} = \Delta(\mathbf{A} \times \Theta_{-i})$ and $\Delta_1 = \prod_{i \in I} \Delta_{i,1}$. If we append to this game a Θ -based type structure $\langle I, (\Theta_i, \mathcal{E}_i, \eta_i)_{i \in I} \rangle$, we obtain a (static) **Bayesian p-game**. The PGT version of the traditional equilibrium concept for such game does not have to be defined in a special—or *ad hoc*—way, we just have to take the Bayesian equilibrium concept “off the shelves,” keeping in mind that any profile of behavior maps yields a belief hierarchy for each type of each player. To allow for non-deterministic plans, we consider **randomized behavior maps** associating each Harsanyi type τ_i with a mixed action $\varsigma_i(\tau_i) = (\varsigma_i(\tau_i)(a_i))_{a_i \in A_i} \in \Delta(A_i)$. Exogenous beliefs are then combined with a given profile $\varsigma = (\varsigma_i : \mathcal{T}_i \rightarrow \Delta(A_i))_{i \in I}$ of randomized behavior maps to obtain the hierarchical endogenous beliefs of each type by means of a generalized pushforward formula. In particular, the first-order belief of any Harsanyi type τ_i is:

$$\alpha_i(\tau_i, \varsigma)(a_i, \mathbf{a}_{-i}, \theta_{-i}) = \varsigma_i(\tau_i)(a_i) \times \left(\sum_{\mathbf{e}_{-i} \in \mathcal{E}_{-i}} \varphi_i(\tau_i)(\theta_{-i}, \mathbf{e}_{-i}) \prod_{j \in I \setminus \{i\}} \varsigma_j(\theta_j, e_j)(a_j) \right) \quad (14)$$

for all $a_i \in A_i$, $\mathbf{a}_{-i} \in \mathbf{A}_{-i}$, $\theta_{-i} \in \Theta_{-i}$. With this, ς is a (randomized) Bayesian equilibrium if, for every $i \in I$ and $(\theta_i, \varepsilon_i) \in \mathcal{T}_i$,

$$\text{supp} \varsigma_i(\theta_i, \varepsilon_i) \subseteq \arg \max_{a_i \in A_i} \mathbb{E} \left[u_i \left(a_i, \varsigma_{-i}(\tilde{\theta}_{-i}, \tilde{\varepsilon}_{-i}), \theta_i, \tilde{\theta}_{-i}, \alpha_i(\theta_i, \varepsilon_i, \varsigma), \boldsymbol{\alpha}_{-i}(\tilde{\theta}_{-i}, \tilde{\varepsilon}_{-i}, \varsigma) \right), \varphi_i(\theta_i, \varepsilon_i) \right].$$

Remark 2 Under complete information, when we consider the special case in which there is only one epistemic type for each player, we obtain the psychological Nash equilibrium of GP&S.

If instead, we allow for multiple epistemic types, we introduce uncertainty about the beliefs of other players (including their plans). Thus, as noticed by Attanasi et al. (2016, Section 3), the Bayesian equilibrium concept allows to reconcile uncertainty about co-players' beliefs with traditional equilibrium analysis, whereas such uncertainty is notably absent from the extensions of traditional complete-information concepts to PGT of GP&S and B&D.

Let us now move on to dynamic games, the main object of our analysis. When we append a Θ -based type structure to a multistage p-game, we obtain a **Bayesian multistage p-game**. In this case, a (randomized) equilibrium is given by a profile

$$\varsigma = (\varsigma_i : \mathcal{T}_i \rightarrow \Sigma_i)_{i \in I} \in \prod_{i \in I} \Sigma_i^{\mathcal{T}_i}$$

of **behavior strategy maps**. In words, each map $\varsigma_i \in \Sigma_i^{\mathcal{T}_i}$ associates each type τ_i of player i with a behavior strategy, that is

$$\varsigma_i(\tau_i) = (\varsigma_i(\cdot|h; \tau_i))_{h \in H} \in \Sigma_i = \prod_{h \in H} \Delta(\mathcal{A}_i(h)).$$

A behavior map ς_i for i is **strictly positive** whenever each mixed action $\varsigma_i(\cdot|h; \tau_i)$ has full support, that is,

$$\varsigma_i(\mathcal{T}_i) \subseteq \Sigma_i^\circ := \prod_{h \in H} \Delta^\circ(\mathcal{A}_i(h)).$$

A profile of behavior strategy maps $\varsigma = (\varsigma_i)_{i \in I}$ is strictly positive whenever ς_i is strictly positive for each $i \in I$. The space of strictly positive behavior strategy maps for i is denoted by $[\Sigma_i^\circ]^{\mathcal{T}_i}$, while the set of profiles of such maps is denoted by $\prod_{j \in I_c} [\Sigma_j^\circ]^{\mathcal{T}_j}$.

Given $(\varphi_i)_{i \in I}$ (exogenous) and $(\varsigma_i)_{i \in I}$ (endogenous), if $(\varsigma_i)_{i \in I}$ is *strictly positive* then, for every $i \in I$, it is possible to derive first- and second-order CPSs for each Harsanyi type $\tau_i = (\theta_i, \varepsilon_i)$, respectively denoted by $\alpha_i(\tau_i, \varsigma)$ and $\beta_i(\tau_i, \varsigma)$, by means of Bayesian updating and a multistage extension of Eq. (14) (we omit the details). Formally, fix a strictly positive profile $\varsigma = (\varsigma_i)_{i \in I}$. There exists a **first-order belief map** $\alpha_i(\varsigma) : \mathcal{T}_i \rightarrow \Delta_{i,1}$, which associates each type τ_i with a corresponding first-order CPS $\alpha_i(\tau_i, \varsigma) = (\alpha_i(\tau_i, \varsigma)(\cdot|h_i))_{h_i \in H_i} \in \Delta_{i,1}$. Given the profile of first-order belief maps $(\alpha_i(\varsigma) : \mathcal{T}_i \rightarrow \Delta_{i,1})_{i \in I}$, for each $i \in I$, we can derive the **second-order belief map** $\beta_i(\varsigma) : \mathcal{T}_i \rightarrow \Delta_{i,2}$. By construction, $\alpha_i(\varsigma)$ is the marginal of $\beta_i(\varsigma)$.

A profile of behavior and second-order belief maps

$$(\varsigma, \beta) = ((\varsigma_i(\tau_i), \beta_i(\tau_i))_{\tau_i \in \mathcal{T}_i})_{i \in I} \in \prod_{i \in I} (\Sigma_i \times \Delta_{i,2})^{\mathcal{T}_i}$$

is called an **assessment**. The following definition adapts K&W's SE concept to assessments in Bayesian multistage p-games. Their sequential rationality condition is replaced by rational planning, which is equivalent to sequential rationality in standard games.⁶⁰ Their consistency

⁶⁰Since utility functions in standard games are not belief-dependent, they necessarily satisfy own-plan independence, hence the equivalence. See Section 7.1.

condition (essentially a very strong independence condition with a “trembling-hand” interpretation) is here adapted to account for higher-order beliefs. This condition implies that observed deviations from the equilibrium path are interpreted as unintentional moves and that they do not lead to the expectation of further deviations.

Definition 13 *Assessment* (ς, β) is a **BSE** of the Bayesian multistage p -game determined by Θ -based type structure $\langle I, (\Theta_i, \mathcal{E}_i, \eta_i)_{i \in I} \rangle$ if

1. **Rational Planning:** For all $i \in I$ and $\tau_i \in \mathcal{T}_i$,

$$\forall h \in H, \text{supp}_{\varsigma_i}(\cdot | h; \tau_i) \subseteq r_{i,h}(\theta_i(\tau_i), \beta_i(\tau_i));$$

2. **K&W-Consistency:** there exists a sequence $(\varsigma_n)_{n \in \mathbb{N}}$ of profiles of strictly positive behavior strategy maps converging to ς and such that, for all $i \in I$ and $\tau_i \in \mathcal{T}_i$,

$$\beta_i(\varsigma_n)(\tau_i) \longrightarrow \beta_i(\tau_i).$$

Unlike one-stage games, in Bayesian multistage p -games (conditional) beliefs cannot be directly derived from behavior strategy maps because some histories are reached with probability 0 in equilibrium, i.e., they are off the equilibrium path. This is not much different from the SE concept for standard games, where off-equilibrium-path beliefs about co-players’ types (or past actions, in games with imperfectly observable actions) are not pinned down by equilibrium strategies *via* Bayes rule. For this reason, in both cases we look for equilibrium assessments rather than just equilibrium strategies. Yet, the spirit of our previous observation still applies: Using the off-the-shelf Bayesian equilibrium concept allows to avoid definitions specifically tailored to p -games. Indeed, hierarchies of initial beliefs can be derived from the behavior strategy maps, which settles the question when only initial beliefs enter in the utility functions (as assumed by GP&S). Furthermore, conditional beliefs can be partially pinned down through an approximation by a sequence of strictly positive behavior strategy maps. The following theorem shows that the BSE concept is not empty.

Result 2 *Every Bayesian multistage p -game with continuous utility functions has at least one BSE.*⁶¹

The most important feature of *BSE* is that it *allows for non-trivial updating about the beliefs of others, including their plans; hence, it allows for non-trivial updating of beliefs about intentions*. This is true also under complete information, as long as the type structure contains multiple epistemic types, as noted by BD&S for a special case of BSE (see point 4 of Remark 3). By contrast, the earlier notions of SE for multistage games with complete information imply that players’ higher-order (point) beliefs are correct and cannot change, on or off the equilibrium path (see GP&S, B&D, and BD&S). This implies that such notions, unlike BSE, do not allow for non-trivial updating about co-players’ intentions. In the following set of remarks, we call **naive** a type structure with only one epistemic type for each player (each \mathcal{E}_i is a singleton); a type structure with only one type for each player is called **trivial**. Thus, a type structure is trivial if it is naive and there is complete information, i.e., Θ is a singleton.

⁶¹This can be shown by adapting the proof of existence of sequential equilibria of B&D. See also BD&S.

Remark 3 With a trivial type structure, $\Delta_{i,1} \subseteq \Delta^{H_i}(Z)$, and in each BSE players share the same first-order beliefs conditional on each history $h \in H$.

Also, recall that we maintain the simplifying assumption of observable actions.

Result 3 BSE is equivalent to known solution concepts in special cases:

1. In standard Bayesian multistage games with complete information, if the type structure is naive—hence trivial—BSE is equivalent to subgame perfect equilibrium (this, of course, relies on our maintained assumption of observable actions, otherwise BSE would be a refinement of subgame perfection).
2. In standard Bayesian multistage games, BSE is equivalent to K&W’s SE.
3. In Bayesian multistage p -games with complete information, if the type structure is naive—hence trivial—BSE is equivalent to the SE concept of BD&S.
4. In Bayesian multistage p -games with complete information, if the type structure is such that each player’s beliefs about others are type-independent (hence, only his plan depends on his type) BSE is equivalent to the “polymorphic” SE concept of BD&S.

Next we illustrate the SE concept obtained as a special case of BSE with a trivial type structure (see point 3 of Remark 3) and show how it differs from the SE concept of GP&S and B&D.

SE in the Trust Minigame with guilt aversion and complete information Consider the Trust Minigame with guilt aversion depicted in Fig. 5, and assume that θ_B is common knowledge, i.e., there is *complete information*. We analyze BSEs when the type structure is naive, hence *trivial* (because there is complete information). Since there is only one type for each player, a BSE is just a profile (pair) of behavior strategies (σ_A, σ_B) and a corresponding profile of second-order beliefs (β_A, β_B) . Furthermore, the correct-conjectures condition implies that each player’s (first-order) belief about the co-player’s behavior coincides with the co-player’s behavior strategy (Remark 3). Thus, in equilibrium, $\alpha_A^{In} = \alpha_B^{In}$ and $\alpha_A^{Sh} = \alpha_B^{Sh}$. Of course, the set of BSEs depends on the commonly known parameter θ_B . Neglecting knife-edge cases for simplicity, there are three possible situations:

- *Low guilt aversion:* $\theta_B < 1$. In this case, the best response of Bob is to *Take*, independently of β_B . Therefore, Ann predicts that Bob would *Take* if she trusts him and we obtain the pure equilibrium strategies (Out, Tk) . The corresponding second-order beliefs are degenerate. For example,

$$\beta_B \left(\left[z = Out, \tilde{\alpha}_A^{In} = \tilde{\alpha}_A^{Sh} = 0 \right] | \emptyset \right) = 1,$$

and so on.⁶² Therefore, $E[D_A, \beta_B | \emptyset] = 0$.

⁶²Since equilibrium assessments are consistent, first-order beliefs can either be derived from the strategies, or from the second-order beliefs.

- *Intermediate guilt aversion:* $1 < \theta_B < 2$. As explained in Section 7.5, if $\theta_B > 1$, Bob's best reply is to *Share* if $E[D_A, \beta_B | In] > 2/\theta_B$. Therefore, there are two pure BSEs, (i) (Out, Tk) with $\alpha_A^{In} = 0$, $\alpha_B^{Sh} = 0$ and, by correct conjectures, $E[D_A, \beta_B | In] = 1 < 2/\theta_B$, and (ii) (In, Sh) with $\alpha_A^{In} = 1$, $\alpha_B^{Sh} = 1$ and $E[D_A, \beta_B | In] = 2 > 2/\theta_B$. Next, consider mixed equilibria in which Bob's plan is not deterministic. The necessary indifference condition is $E[D_A, \beta_B | In] = 2/\theta_B$. There are three subcases. Note that the consistency condition holds.
 - If $\alpha_B^{Sh} = \alpha_A^{Sh} < 1/2$ then Ann best responds with *Out*, that is, $\alpha_A^{In} = 0$. Since Bob is certain of Ann's plan, we have $E[D_A, \beta_B | In] = 1 < 2/\theta_B$, which contradicts the indifference condition.
 - If $\alpha_B^{Sh} = \alpha_A^{Sh} = 1/2$ then Ann is indifferent, that is, $\alpha_A^{In} \in [0, 1]$. Since Bob is certain of Ann's first-order belief, we have again $E[D_A, \beta_B | In] = 1 < 2/\theta_B$, which contradicts the indifference condition.
 - If $\alpha_B^{Sh} = \alpha_A^{Sh} > 1/2$ then Ann's best reply is *In*, that is, $\alpha_A^{In} = 1$. Given that Bob is certain of Ann's first-order belief, $E[D_A, \beta_B | In] = 2\alpha_B^{Sh} > 1$. Finally, the indifference condition gives $\alpha_B^{Sh} = 1/\theta_B \in (1/2, 1)$; hence, there is a mixed BSE where Ann goes *In* and Bob *Shares* with probability $1/\theta_B$.
- *High guilt aversion:* $\theta_B > 2$. In this case there is only one BSE, (In, Sh) . To see this, note that Ann has the option of securing a payoff of \$1; hence, in equilibrium she must expect to get at least \$1. Since Bob's second-order beliefs are correct, he is certain—both ex ante and upon observing *In*—that the disappointment caused by *Take* is at least 1. Hence, he wants to *Share* if $2 > 4 - \theta_B$, that is, $\theta_B > 2$, the case we are considering. Thus, Ann predicts that Bob would *Share* and chooses to trust.

Note the difference with the equilibrium analysis of B&D, where (Out, Tk) is an equilibrium for every θ_B . Formally, B&D consider a different psychological utility for Bob at terminal history (In, Tk) . But the deeper reason is that B&D model explicitly only players' beliefs about co-players. As they acknowledge, they implicitly assume that players' plans necessarily coincide with their behavior, which implies that if Bob observes *In* he is certain that Ann planned to trust. But Bob's second-order belief about Ann's initial expectation that he would *Share* is correct, hence the same at each node. This implies that, if Bob is initially certain that Ann expects him to *Take* and nonetheless he observes *In*, he becomes certain that Ann expects to get \$0, hence he wants to *Take*.

8 Concluding remarks

Reciprocity, emotions, concerns for others' opinions, and self-esteem are belief-dependent forms of motivation. Theoretical modelling leads to p-games. In this contribution, we have shown how. Our aim has been to help scholars who wish to do applied economics and to explore how those sentiments shape outcomes. We also scrutinized a variety of (old & new) PGT solution concepts. We illustrated our approach starting from the Trust and the Ultimatum Minigames, two game forms where the credibility of promises (Trust) and threats (Ultimatum) is crucial. We showed

how belief-dependent motivations, such as guilt aversion (Trust) and anger (Ultimatum) enhance the credibility of promises and threats.

Our approach is broad enough to address many concerns. Yet, there are themes we have not explored either to simplify the analysis, or because there are some technical problems to address, or because there are conceptual problems to address. We close this paper by listing some of these issues, left for future work.

More general game forms We assumed for simplicity that the game form has finitely many perfectly observable actions, but we can extend the analysis to games with imperfectly observable actions, and also allow, to some extent, for a continuum of actions and personal traits.

Causal effects of beliefs We analyzed how beliefs affect material outcomes (consequences) indirectly *via* belief-dependent motivations (which drive agents’ behavior). Yet, beliefs may affect outcomes more directly *via* induced emotions, e.g., when anxiety affects performance (e.g., Rauh & Seccia 2006), or anger affects strength (Gneezy & Imas 2014). On the one hand, though we didn’t pursue the theme, such effects can be captured using the reduced-form psy-utility functions u_i of Section 6. On the other hand, our framework, as is, does not allow to represent the impact of (belief-dependent) emotions on cognitive abilities, which is also acknowledged in the aforementioned papers.

SCE in population games and games with chance moves Our simplified analysis of SCE excludes the objective randomness of population games and more generally of games with chance moves. Applications often feature such randomness, which can be accounted for by extending to p-games the “mixed” SCE concept of Fudenberg & Levine (1993).⁶³

Signals about emotions In face-to-face (or otherwise non-anonymous) interaction, agents observe signals, such as facial cues, about personal traits (e.g., trait-anger) and emotions (e.g., state-anger) of others (see van Leeuwen et al. 2018). Our analysis can be extended to include such signals. This entails addressing technical issues concerning signals about continuous variables.

Game-dependent preferences and personal traits We emphasized a distinction between the belief-dependent preferences that can be modeled without the game form and those that we can only model with reference to a game form. While we find this distinction useful, we also think it reflects a limitation of our analysis and, to the best of our knowledge, of current PGT. The issue with game-dependent preferences is that they involve key elements related to the causal structure of the decision problem, like the value of the best unchosen alternative in the case of regret (see the illustrative example in Section 6), and we recover them from the game form. Yet, preferences are personal traits presumably shaped in our evolutionary past. Assuming that the archive of possible game forms is somehow encoded in our neural circuits stretches credulity. Thus, there should be a way to describe such preferences without direct reference to game forms. This could perhaps be achieved by introducing “interface variables” in the model, that is, variables that affect psychological states whose meaning we can understand with no reference to a specific decision

⁶³See also Battigalli et al. (1992) and Battigalli et al. (2015).

problem, but whose value is determined by the decision problem itself. Again, the value of the best alternative in a model of regret may be an example.

References

- [1] Aina, C., P. Battigalli, and A. Gamba (2018): “Frustration and Anger in the Ultimatum Game: An Experiment,” IGIER W.P. 621, Bocconi University.
- [2] Attanasi, G., P. Battigalli, and E. Manzoni (2016): “Incomplete Information Models of Guilt Aversion in the Trust Game,” *Manage. Sci.*, 62, 648-667.
- [3] Attanasi, G., P. Battigalli, E. Manzoni, and R. Nagel (2018): “Belief-Dependent Preferences and Reputation: Experimental Analysis of a Repeated Trust Game,” *J. Econ. Behav. Org.*, <https://doi.org/10.1016/j.jebo.2018.12.002>.
- [4] Attanasi G., P. Battigalli, and R. Nagel (2013): “Disclosure of Belief-Dependent Preferences in the Trust Game,” IGIER W.P. 506, Bocconi University.
- [5] Aumann, R.J. (1974): “Subjectivity and Correlation in Randomized Strategies,” *J. Math. Econ.*, 1, 67-96.
- [6] Battigalli, P., S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2015): “Self-Confirming Equilibrium and Model Uncertainty,” *Am. Econ. Rev.*, 105, 646-677.
- [7] Battigalli P., G. Charness, and M. Dufwenberg (2013): “Deception: The Role of Guilt,” *J. Econ. Behav. Org.*, 93, 227-232.
- [8] Battigalli P., R. Corrao, and F. Sanna (2019): “Epistemic Game Theory without Types Structures: An Application to Psychological Games,” IGIER W.P. 641, Bocconi University.
- [9] Battigalli, P., and N. De Vito (2018): “Beliefs, Plans, and Perceived Intentions in Dynamic Games,” IGIER W.P. 629.
- [10] Battigalli, P., A. Di Tillio, and D. Samet (2013): “Strategies and Interactive Beliefs in Dynamic Games,” in *Advances in Economics and Econometrics*, ed. by D. Acemoglu, M. Arellano, and E. Dekel. Cambridge, UK: Cambridge University Press, 391-422.
- [11] Battigalli P., and M. Dufwenberg (2007): “Guilt in Games,” *Am. Econ. Rev. (P&P)*, 97, 170-176.
- [12] Battigalli P., and M. Dufwenberg (2009): “Dynamic Psychological Games,” *J. Econ. Theory*, 144, 1-35.
- [13] Battigalli P., M. Dufwenberg, and A. Smith (2015): “Frustration and Anger in Games,” IGIER W.P. 539, Bocconi University.
- [14] Battigalli, P., M. Gilli and M.C. Molinari (1992): “Learning and Convergence to Equilibrium in Repeated Strategic Interaction,” *Res. Econ.*, 46, 335-378.

- [15] Battigalli, P., and D. Guaitoli (1997): “Conjectural Equilibria and Rationalizability in a Game with Incomplete Information,” in *Decisions, Games and Markets*, ed. by P. Battigalli, A. Montesano, and F. Panunzi. Dordrecht: Kluwer Academic Publishers, 97-124.
- [16] Battigalli, P., and M. Siniscalchi (1999): “Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games,” *J. Econ. Theory*, 88, 188-230.
- [17] Battigalli, P., and M. Siniscalchi (2002): “Strong Belief and Forward Induction Reasoning,” *J. Econ. Theory*, 106, 356-391.
- [18] Ben-Porath, E. (1997): “Rationality, Nash Equilibrium, and Backward Induction in Perfect Information Games,” *Rev. Econ. Stud.*, 64, 23-46.
- [19] Bernheim, D. (1994): “A Theory of Conformity,” *J. Pol. Econ.*, 102, 841-77.
- [20] Bjorndahl, A., J. Halpern, R. Pass (2013): “Language-based Games,” *Proceedings of Fourteenth Conference on Theoretical Aspects of Rationality and Knowledge*.
- [21] Caplin, A., and J. Leahy (2001): “Psychological Expected Utility Theory and Anticipatory Feelings,” *Q. J. Econ.*, 116, 55-79.
- [22] Caplin, A., and J. Leahy (2004): “The Supply of Information by a Concerned Expert,” *Econ. J.*, 114, 487-505.
- [23] Charness G., and M. Dufwenberg (2006): “Promises and Partnership,” *Econometrica* 74, 1579-1601.
- [24] Dufwenberg, M., (2002): “Marital Investment, Time Consistency and Emotions,” *J. Econ. Behav. Org.*, 48, 57-69.
- [25] Dufwenberg, M., and Dufwenberg, M.A. (2018): “Lies in Disguise—A Theoretical Analysis of Cheating,” *J. Econ. Theory*, 175, 248-264.
- [26] Dufwenberg, M., and U. Gneezy (2000): “Measuring Beliefs in an Experimental Lost Wallet Game,” *Games Econ. Behav.*, 30, 163-182.
- [27] Dufwenberg, M., and G. Kirchsteiger (2004): “A Theory of Sequential Reciprocity,” *Games Econ. Behav.*, 47, 268-298.
- [28] Dufwenberg, M., F. Li, and A. Smith (2018a): “Promises and Punishment,” unpublished manuscript.
- [29] Dufwenberg, M., F. Li, and A. Smith (2018b): “Threats,” unpublished manuscript.
- [30] Dufwenberg, M., and K. Nordblom (2018): “Tax Evasion with a Conscience,” unpublished manuscript.
- [31] Ellingsen, T., and M. Johannesson (2008): “Pride and Prejudice: The Human Side of Incentive Theory,” *Am. Econ. Rev.*, 98, 990-1008.

- [32] Elster, J. (1998): "Emotions and Economic Theory," *J. Econ. Lit.*, 36, 47-74.
- [33] Ely, J., Frankel A., and E. Kamenica (2015) "Suspense and Surprise," *J. Polit. Econ.*, 123, 215-260.
- [34] Falk, A., and U. Fischbacher (2006): "A Theory of Reciprocity," *Games Econ. Behav.*, 54, 293-315.
- [35] Fudenberg, D., and D.K. Levine (1993): "Self-Confirming Equilibrium," *Econometrica*, 61, 523-545.
- [36] Geanakoplos J., D. Pearce, and E. Stacchetti (1989): "Psychological Games and Sequential Rationality," *Games Econ. Behav.*, 1, 60-79.
- [37] Gneezy, U., and A. Imas (2014): "Materazzi Effect and the Strategic Use of Anger in Competitive Interactions," *Proceedings of the Nat'l Academy of Sciences* 111, 1334-1337.
- [38] Gneezy, U., A. Kajackaite, and J. Sobel (2018): "Lie Aversion and the Size of a Lie," *Am. Econ. Rev.*, 108, 419-453.
- [39] Gul, F., and W. Pesendorfer (2016): "Interdependent Preference Models as a Theory of Intentions," *J. Econ. Theory*, 165, 179-208..
- [40] Harsanyi, J. (1967-68): "Games of Incomplete Information Played by Bayesian Players. Parts I, II, III," *Manage. Sci.*, 14, 159-182, 320-334, 486-502.
- [41] Huang, P.H., and H.M. Wu (1994): "More Order without More Law: A Theory of Social Norms and Organizational Cultures," *J. Law Econ. Organ.*, 12, 390-406.
- [42] Jagau, S., and A. Perea (2018): "Common Belief in Rationality in Psychological Games," Epicenter W.P. 10.
- [43] Jang, D., A. Patel, and M. Dufwenberg (2018), "Agreements with Reciprocity: Co-Financing and MOUs," *Games Econ. Behav.*, 111, 85-99.
- [44] Kahneman, D. (1994): "New Challenges to the Rationality Assumption," *J. Inst., and Theo. Econ.*, 150, 18-36.
- [45] Kahneman, D., P. Wakker, and R. Sarin (1997): "Back to Bentham? Explorations of Experienced Utility," *Q. J. Econ.*, 112, 375-405.
- [46] Khalmetski, K., A. Ockenfels, and P. Werner (2015): "Surprising gifts: Theory and laboratory evidence," *J. Econ. Theory*, 159, 163-208.
- [47] Koszegi, B., and M. Rabin (2006): "A Model of Reference-dependent Preferences," *Q. J. Econ.*, 121, 1133-1166.
- [48] Koszegi, B., and M. Rabin (2009): "Reference-dependent Consumption Plans," *Am. Econ. Rev.*, 99, 909-936.

- [49] Kreps, D., and R. Wilson (1982): “Sequential Equilibria,” *Econometrica*, 50, 863-894.
- [50] Kuhnen, C.M., and A. Tymula (2012): “Feedback, Self-Esteem, and Performance in Organizations,” *Manage. Sci.*, 58, 94-113.
- [51] Laibson, D. (1997): “Golden Eggs and Hyperbolic Discounting,” *Quart. J. Econ.*, 112, 443-477.
- [52] van Leeuwen, B., C. Noussair, T. Offerman, S. Suetens, M. van Veelen, and J. van de Ven (2018): “Predictably Angry – Facial Cues Provide a Credible Signal of Destructive Behavior,” *Manage.Sci.*, 64, 2973-3468.
- [53] Levine, D.K. (1998): “Modeling Altruism and Spitefulness in Game Experiments,” *Rev. Econ. Dyn.*, 1, 593-622.
- [54] Loewenstein, G. (2007): *Exotic Preferences: Behavioral Economics and Human Motivation*. Oxford, Oxford University Press.
- [55] Loewenstein, G., C. Hsee, E. Weber, and N. Welch (2001): “Risk as Feelings,” *Psych. Bull.*, 127, 267-286.
- [56] Osborne, M., and A. Rubinstein (1994): *A Course in Game Theory*. Cambridge MA: MIT Press.
- [57] Pearce, D. (1984): “Rationalizable Strategic Behavior and the Problem of Perfection,” *Econometrica*, 52, 1029-1050.
- [58] Perea, A. (2014): “Belief in the Opponents’ Future Rationality,” *Games Econ. Behav.*, 83, 231-254.
- [59] Phelps, E., and R. Pollak (1968): “On Second-Best National Saving and Game-Equilibrium Growth,” *Rev. Econ. Stud.*, 35, 185-199.
- [60] Rabin, M. (1993): “Incorporating Fairness into Game Theory and Economics,” *Am. Econ. Rev.*, 83, 1281-1302.
- [61] Rauh, M., and G. Seccia (2006): “Anxiety and Performance: a Learning-by-Doing Model,” *Int. Econ. Rev.*, 47, 583-609.
- [62] Rothmann, S., and E.P. Coetzer (2003): “The Big Five Personality Dimensions and Job Performance,” *SA J. Industrial Psychology*, 29, 68-74.
- [63] Rubinstein, A., and A. Wolinsky (1994): “Rationalizable Conjectural Equilibrium: Between Nash and Rationalizability,” *Games Econ. Behav.*, 6, 299-311.
- [64] Selten, R. (1975): “Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Games,” *Intl. J. Game Theory*, 4, 25-55.
- [65] Tangney, J.P. (1995): “Recent Advances in the Empirical Study of Shame and Guilt,” *American Behavioral Scientist*, 38, 1132-1145.