

# Strategic Thinking and Value of Reasoning: Theory and Applications to Five ‘Little Treasures’ of Game Theory\*

Larbi Alaoui<sup>†</sup>  
Universitat Pompeu Fabra

Antonio Penta<sup>‡</sup>  
Dept. of Econ., UW-Madison

March 12, 2013

## Abstract

We develop and analyze an axiomatic model of strategic thinking in games, and demonstrate that it accommodates well-known empirical puzzles that are inconsistent with standard game theory. We model the reasoning process as a Turing machine, thereby capturing the stepwise nature of the procedure. Each ‘cognitive state’ of the machine is a complete description of the agent’s state of mind at a stage of the reasoning process, and includes his current understanding of the game and his attitude towards thinking more. We take axioms over the reasoning about games within each cognitive state. These axioms restrict the scope of the reasoning process to understanding the opponent’s behavior. They require that the value of reasoning be purely instrumental to improving the agent’s choice, and relate the suitability of the reasoning process to the game at hand. We then derive representation theorems which take the form of a cost-benefit analysis. Players behave as if they weigh the instrumental value of thinking deeper against the cost. To compare behavior across different games, we introduce a notion of cognitive equivalence between games. Games within the same cognitive equivalence class induce the same cost of reasoning function and only differ in the incentives to reason, where incentives are related to the difference in payoffs across actions. We further enrich our model and allow the agent to account for the opponent’s own cost-benefit reasoning procedure. Lastly, we apply our model to Goeree and Holt’s (2001) well-known ‘little treasures’ games. We show that the model is consistent with all the ‘treasures’ that fall within the domain of our theory. We perform a single-parameter calibration exercise using stringent restrictions on the model’s degrees of freedom, and show that the model closely matches the experimental data.

**Keywords:** cognitive cost – depth of reasoning – introspection – level-k reasoning – little treasures – strategic thinking

**JEL Codes:**C72; C92; D80; D83.

---

\*The authors gratefully acknowledge financial support respectively from the Spanish Ministry of Science and Innovation under project ECO2011-25295 (Alaoui) and the Research Grant provided by the Graduate School at the University of Wisconsin-Madison (Penta).

<sup>†</sup>Universitat Pompeu Fabra and Barcelona GSE. E-mail: larbi.alaoui@upf.edu

<sup>‡</sup>University of Wisconsin-Madison. E-mail: apenta@ssc.wisc.edu

# 1 Introduction

Individual behavior systematically deviates from the predictions of standard game theory in a large number of strategic settings. Since the seminal work of Nagel (1995) and Stahl and Wilson (1994, 1995), studies of ‘initial responses’ have become a major path of inquiry to uncover the cognitive processes that underlie individuals’ strategic choices. This literature shows that, when playing games without clear precedents, individuals typically do not appear to perform the kind of fixed point reasoning entailed by classical equilibrium concepts. Rather, individuals’ choices reveal distinct patterns that suggest that the thinking process occurs in steps.<sup>1</sup> In an influential study of initial responses, however, Goeree and Holt (2001) find that individual play is consistent with equilibrium predictions for some parameters, but inconsistent for others. The findings of Goeree and Holt do not bring to light a clear pattern of reasoning, but the intuitive appeal of these results suggests that they are the consequence of a fundamental underlying reasoning process. Our aim in this paper is to provide a tractable unifying framework which makes explicit the reasoning procedure in strategic settings. We analyze the foundational properties that shape this process, and demonstrate that, using the same minimal parametric assumptions throughout, this model is highly consistent with well-known empirical results and with Goeree and Holt’s ‘intuitive contradictions’. We then illustrate that the mechanisms behind these results may serve to better understand and isolate the forces behind the informal intuition.

We view reasoning as stemming from a costly introspection procedure in which it is *as if*, at each step of the reasoning process, the player weighs the value of thinking deeper against a cost. We further relate the ‘value of thinking’ to the payoff structure of the game, and endogenize the number of steps that the agent performs, or his ‘depth of reasoning’. Thus, in our approach, incentives not only affect individual choices in the traditional sense, but they also shape the cognitive process itself.<sup>2</sup> As shown in Alaoui and Penta (2013), acknowledging the interaction between cognition and incentives in games is empirically relevant, and may be instrumental to a better understanding of existing models of strategic thinking.<sup>3</sup> Furthermore, from a methodological viewpoint, this approach has the advantage of bridging the study of strategic thinking with standard economics concepts, opening new directions of research.

From a naive perspective, a cost-benefit tradeoff appears natural and within the core of standard economic theory. But reasoning about opponents’ choices is not a conventional domain of analysis, and it is distinct from other realms of bounded rationality, such as complexity and costly contemplation over processing information. It is then not obvious how the value of reasoning should be perceived, which elemental properties should hold, or how the tradeoff with the costs should be characterized. To address these questions, here we follow the classical decision theory methodology. That is, we take axioms over the primitive preferences to aid our understanding of

---

<sup>1</sup>For a recent survey on the empirical and theoretical literature on initial responses see Crawford, Costa-Gomes and Iriberry (2012). For studies that focus more directly on the cognitive process itself, see Agranov, Caplin and Tergiman (2012), or the recent works by Bhatt and Camerer (2005), Coricelli and Nagel (2009), and Bhatt, Lohrenz, Camerer and Montague (2010), which use fMRI methods finding further support for level-k models. For recent theoretical work inspired by these ideas, see Strzalecki (2010), Kets (2012) and Kneeland (2013).

<sup>2</sup>It is a well known theme in the Economics of Education literature that incentives may affect standard measures of cognitive and non cognitive abilities. For a recent survey of the vast literature that combines classical economic notions such as comparative advantage and incentives with measurement of cognitive abilities and psychological traits, see for instance, Almlund et al. (2011).

<sup>3</sup>In Alaoui and Penta (2013), we show that the cost-benefit approach delivers a rich set of predictions while maintaining a general structure. We test these predictions experimentally, and find that our model provides a unifying framework that accommodates the empirical evidence. Existing models that take individuals’ depth of reasoning as exogenous cannot jointly account for these findings.

the deeper forces behind the cost-benefit representation. Our domain of preferences is atypical, however. Since the agent’s fundamental decision problem is inseparable from the strategic context, this leads us to directly consider preferences over games. In this sense, we take a step towards providing a framework for modeling individual reasoning process in strategic settings.

We begin our analysis with the notion of ‘cognitive state’. To make explicit the underlying premise of our approach, we take a cognitive state to be a complete description of the agent’s state of mind at some stage of a reasoning process. We do not impose restrictions on the richness of these states, and therefore any stepwise reasoning process can be modeled with no essential loss of generality as a Turing machine defined over this set of states. At a minimum, however, a cognitive state includes the agent’s current understanding of the game and his attitude towards further pursuing the reasoning process, which we model as a preference system over lotteries of games cross the choice to perform an additional round of reasoning. We then impose standard continuity and monotonicity axioms, as well as others that are suited to this problem. First, we impose axioms that restrict the scope of the problem by removing orthogonal considerations. In particular, one axiom requires that the value of reasoning is purely instrumental and comes only from the agent’s potential payoff gain. Furthermore, since our focus is on understanding reasoning about the opponents, we also ignore issues of computational complexity in determining one’s own best responses. We focus instead on the cognitive complexity of refining one’s understanding of the opponents’ behavior. Lastly, our ‘unimprovability’ axiom effectively requires that the decision maker’s reasoning process is apt to the game he is playing. We discuss these novel axioms in detail, as they serve to elucidate the nature of the reasoning process and of the ‘as-if’ interpretation behind the cost-benefit analysis.

Our axioms restrict the agent’s preferences within each mental state, but we impose no restriction across states. Our representation therefore concerns the determinants of the value of reasoning and the agent’s ‘choice’ of whether or not to reason further. But it does not restrict how a specific reasoning process, or machine, transitions from one state to another, or how it relates to other reasoning processes. Our approach is therefore consistent with different specifications of the reasoning process itself. Given an arbitrary rule that assigns games to reasoning processes, we define a notion of cognitive similarity between games. This notion allows us to define the concept of comparability of the reasoning process across games. Games within the same cognitive equivalence class induce the same cost of reasoning function and only differ in the incentives to reason, which allows for comparative statics on the depth of reasoning as the payoffs are changed. For instance, it may seem plausible that two games that are identical up to positive affine transformation are equally ‘difficult’ to reason about, but that two games with different best responses are not. Our model accommodates this view as well as other examples of cognitive partitions.

We then apply our model to analyzing the findings of Goeree and Holt’s (2001) five static ‘little treasures’. These intuitive results are difficult to reconcile with standard game theory and they cannot be explained by existing models of bounded rationality, but we show that they are well within our model. In particular, we take the most stringent parametric restrictions of our model, and append it with a transition function across mental states consistent with standard level- $k$  models. Maintaining the same assumptions throughout, we perform a calibration exercise. Our model does not only fit the qualitative results, it is also highly consistent with the quantitative findings under these inflexible parameters. Hence, in addition to serving as a unifying theoretical framework, this model accommodates well-known empirical results from a wide range of studies

and provides a mechanism for the intuition underlying many known experiments.

The rest of the paper is structured as follows. Section 2 introduces the general framework and provides a brief overview of the representation theorem. Section 3 provides the axioms and main theorems for the decision maker’s reasoning over the game, and Section 4 further adds the component of the agent’s beliefs over the opponent’s reasoning. Section 5 then demonstrates that our model is highly consistent with the static ‘treasures’, and Section 6 concludes. All proofs are in the Appendix.

## 2 The Model of Reasoning

Consider an individual who, in attempting to predict his opponent’s behavior, proceeds in steps instead of following an equilibrium-like fixed-point reasoning procedure. For instance, consider the process of iteratively eliminating dominated strategies. In the first step, player  $i$  deletes options that a rational player  $j$  would never play. In the second step, player  $i$  further deletes player  $j$ ’s strategies that are only justified by conjectures that involve dominated strategies for player  $i$ , and so forth. Alternatively, as posited by models of ‘level- $k$  reasoning’, in the first step player  $i$  believes that player  $j$  would play a specific action (possibly a mixture), and best responds to that. In the second step, player  $i$  takes into account that player  $j$  may anticipate  $i$ ’s first round, and best respond to  $j$ ’s response to  $i$ ’s first round, and so on.

The common feature of these two examples is that the player follows a stepwise reasoning procedure. This is a key aspect of our general framework. Given such a stepwise nature, we endogenize the point, if any, at which the reasoning process may be interrupted. While it is customary to think of classical game theory as the domain of unboundedly rational players who can draw all the conclusions from a given line of reasoning, cognitively bounded individuals may interrupt the reasoning process earlier. Hence, individuals with different cognitive abilities may have different ‘depth of reasoning’ in a given game. But it may also be that an agent’s depth of reasoning may vary across different strategic situations. Our general aim is to understand how the ‘depth of reasoning’ varies as a function of the individuals’ cognitive abilities and the payoff structure of the game.

The distinctive feature of our model is that the depth of reasoning results from an ‘as if’ cost-benefit analysis. Intuitively, each step of reasoning has value to the player. For boundedly rational individuals, however, such steps may be costly to conduct, and it would be ‘optimal’ to interrupt the reasoning process if the cost exceeds the ‘value of reasoning’. Formally, we will represent such cost and benefits through functions  $c_i : \mathbb{N} \rightarrow \mathbb{R}_+$  and  $W_i : \mathbb{N} \rightarrow \mathbb{R}_+$ : for each  $k \in \mathbb{N}$ , an individual who has performed  $k - 1$  steps of reasoning will perform the next step if and only if the cost  $c_i(k)$  of performing the  $k$ -th step is no larger than its value,  $W_i(k)$ .

We emphasize that these costs and values are not meant in a literal sense. Rather, we view the ‘depth of reasoning’ itself as an *outcome* of the reasoning process, not a conscious choice of the agent. However, to the extent that individuals’ depth of reasoning varies in a systematic way across different strategic situations, then it is *as if* individuals face a tension between cognitive costs and ‘value of reasoning’, which are related to the payoffs of the underlying strategic situation. Our cost-benefit approach is thus a result, which follows from more primitive assumptions on the reasoning process.

The cost-benefit representation is very convenient, particularly because it enables us to perform

comparative statics exercises on the depth of reasoning as the payoffs of the game are varied. But strategic thinking is not a standard domain for a cost-benefit approach, and there are several non obvious modeling choices that might seem arbitrary without a clear foundation. For instance, it is not clear how the cost and value of reasoning functions should be related to the game payoffs. Moreover, the comparative statics exercise of increasing the value of reasoning  $W_i$  through changing the game payoffs, while holding  $c_i$  fixed, requires a notion of which games are equally ‘difficult’ from a cognitive viewpoint, so that their differences would only affect the incentives to reason, not the cognitive costs. The notion of ‘value of reasoning’ itself is more subtle than in other apparently related contexts (such as, for instance, the value of information) because it should capture the idea, inherent to the very notion of bounded rationality, that the agent does not know (or is not aware of) whatever he has not (yet) thought about. That is, for each  $k \in \mathbb{N}$ ,  $W_i(k)$  represents the value that player  $i$  assigns to performing the  $k$ -th step of reasoning, *given* that he has performed  $k - 1$  rounds, and this in general need not be consistent with what player  $i$  would actually learn from the  $k$ -th step of reasoning, or with  $W_i(l)$ , for  $l \neq k$ .

In this section we introduce our general framework, and briefly review the main cost-benefit representations that will be used in Section 5 for the applications. The axioms and the derivation of these representations are postponed to Section 3.

## 2.1 Games as ‘Menus of Acts’

Consider a game form  $\mathcal{F} = \langle N, (A_i)_{i \in N} \rangle$ , where  $N$  is a finite set of players, each with a finite action set  $A_i$  (for each  $i \in N$ ), and let  $A = \times_{i \in N} A_i$  and  $A_{-i} = \times_{j \neq i} A_j$ . A game is obtained attaching payoff functions  $u_i : A \rightarrow \mathbb{R}$  to the game form:  $G = (\mathcal{F}, (u_i)_{i \in N})$ . Holding the game form constant, let  $\mathcal{U} = \mathbb{R}^{n|A|}$  denote the set of all payoffs profiles functions, and  $\mathcal{U}_i = \mathbb{R}^{|A|}$  the set of all payoff functions for player  $i$ . Each game is thus identified by a point  $u \in \mathcal{U}$ , and we let  $u_i \in \mathcal{U}_i$  denote the component of  $u$  corresponding to player  $i$ . These payoffs are measured in ‘utils’.

Given payoff function  $u_i \in \mathcal{U}_i$ , each action  $a_i \in A_i$  determines an *act*  $u_i(a_i, \cdot) : A_{-i} \rightarrow \mathbb{R}$  that pays out in utils, as a function of the realization of the opponents’ actions. Each payoff function  $u_i$  can thus be seen as a *menu of acts* for player  $i$ , and each game  $u \in \mathcal{U}$  as a collection of menus, one for each player. For any  $u_i, v_i \in \mathcal{U}_i$ , and for any  $\alpha \in [0, 1]$ , we denote by  $\alpha u_i + (1 - \alpha) v_i \in \mathcal{U}_i$  the payoff function (or menu of acts) that pays  $\alpha u_i(a) + (1 - \alpha) v_i(a)$  for each  $a \in A$ .

We will use the following notation: for any set  $X$ , we let  $\Delta(X)$  denote the set of simple probability measures on  $X$ , and specifically agents’ conjectures in the game are denoted by  $\mu^i \in \Delta(A_{-i})$ . We define player  $i$ ’s best response in game  $u$  as  $BR_i^u(\mu) = \arg \max_{a_i \in A_i} \sum_{a_{-i}} \mu(a_{-i}) u_i(a_i, a_{-i})$ . We sometimes write  $BR_i^u(\alpha)$  for  $\alpha \in \Delta(A)$  instead of  $BR_i^u(\text{marg}_{A_{-i}} \alpha)$ , and  $BR_i^u(a_{-i})$  for the best response to the conjecture concentrated on  $a_{-i} \in A_{-i}$ . We will also adopt the following notation for vectors: for any  $x, y \in \mathbb{R}^n$ , we let  $x \geq y$  denote the case in which  $x$  is weakly larger than  $y$  for each component, but not  $x = y$ ;  $x \geq y$  means  $x \geq y$  or  $x = y$ ; we let  $x \gg y$  denote strict inequality for each component. For  $c \in \mathbb{R}$  and  $x \in \mathbb{R}^n$ , we abuse notation slightly and write  $x + c$  to mean  $x + c \cdot \mathbf{1}$ , where  $\mathbf{1} \in \mathbb{R}^n$ .

## 2.2 The Reasoning Process as a Machine

In the following we will focus on a specific player  $i$ , and model his reasoning process about the games that he may have to play. We let  $\Pi$  denote the set of all possible *reasoning processes*

that this player may follow, depending on the specific context, and we let  $S$  denote the set of all possible ‘mental states’ of this player. Each  $s \in S$  should be interpreted as a full description of the player’s reasoning at some point of some reasoning process. Without loss of generality, we can represent each reasoning process  $\pi \in \Pi$  as a machine  $\pi = \langle s^{\pi,0}, \tau^\pi \rangle$  over the set of states  $S$ , where  $s^{\pi,0} \in S$  is the initial state, and  $\tau^\pi : S \rightarrow S$  is the transition function, which describes how the process moves from a mental state to another. Given  $\pi$ , we denote by  $s^{\pi,k}$  the  $k$ -th state obtained iterating  $\tau^\pi$  from  $s^{\pi,0}$ . That is, for  $k = 1, 2, \dots$ , we let  $s^{\pi,k} = \tau^\pi(s^{\pi,k-1})$ . (The reference to  $\pi$  will be dropped when the relevant reasoning process is clear from the context.) Finally, we let  $S(\pi)$  denote the set of mental states that arise if the agent adopts reasoning process  $\pi$ :  $S(\pi) = \{s \in S : s = s^{\pi,k} \text{ for some } k \in \mathbb{N}\}$ .

For any game  $u \in \mathcal{U}$ , we let  $\pi(u)$  denote the reasoning process that a particular player  $i$  adopts to reason about  $u$ . Hence, each game  $u \in \mathcal{U}$  induces a sequence of states  $\{s^{\pi(u),k}\}_{k=0}^\infty \equiv S(\pi(u)) \subset S$ , which can be thought of as the set of mental states that arise in the reasoning process generated by the game  $u$ . Notice that here we are making the implicit assumption that only the payoffs of the game determine player  $i$ ’s reasoning process. This restriction, however, can be easily relaxed, letting  $\pi$  depend on other characteristics of the ‘context’. This may be an interesting generalization, but we ignore it here.<sup>4</sup>

We assume that, in each mental state  $s \in S$ , player  $i$  has a particular action  $a_i^s$  that he currently regards as ‘the most sophisticated’ as well as a well-defined system of preferences over ‘which game to play’ and whether he would prefer to play it with or without further thinking about the strategic situation. We formalize this as a preference system  $\succsim_s$  over  $\mathcal{U}_i \times \{0, 1\}$ , which will be discussed below.<sup>5</sup> We impose the following condition, which requires that the reasoning process is optimal in a minimal sense at every mental state:

**Condition 1** *We assume that for any  $u \in \mathcal{U}$  and for any  $s \in S(\pi(u))$ , action  $a_i^s$  is non-dominated in  $u$ .*

We will denote a generic element  $(u_i, x) \in \mathcal{U}_i \times \{0, 1\}$  by  $u_i^x$ , and for  $x = 0, 1$  we let  $\succsim_s^x$  denote the restriction of  $\succsim_s$  on  $\mathcal{U}_i \times \{x\}$ . Preferences  $\succsim_s^0$  are best interpreted as preferences over which game to play based on the understanding at state  $s$ . Preferences  $\succsim_s^1$  instead can be interpreted as the preferences that an agent in mental state  $s$  has about which game to play, if his choice was deferred until the agent had performed the next round of reasoning, and the agent played according to his most sophisticated understanding of the game. For instance:  $u_i^0$  corresponds to the decision problem of agent  $i$  is ‘choose an act in  $(u_i(a_i, \cdot))_{a_i \in A_i}$  without any further reasoning’, whereas  $u_i^1$  corresponds to the problem ‘choose an act in  $(u_i(a_i, \cdot))_{a_i \in A_i}$  after having reasoned more about it’;  $u_i^1 \succsim_s v_i^0$  means ‘given his understanding at mental state  $s$ , agent  $i$  would rather choose an act in  $u_i$  after having thought more, than choose an act in  $v_i$  without further thinking’; and so on. These preference systems represent the agent’s attitude towards the reasoning process, including whether or not it should be continued.

<sup>4</sup>We thus identify the payoff structure with the ‘context’ which determines the reasoning process. Because the entire analysis maintains a given player  $i$ , however, the function that assigns games to reasoning processes may be individual specific. We make no assumptions on interpersonal comparisons of the rule that assigns games to reasoning processes.

<sup>5</sup>Given the dependence of  $s$  on  $\pi(u)$ , the fact that such preferences over which game to play only depend on  $u_i$  entails no loss of generality, because the dependence on  $u_{-i}$  is embedded in the dependence of the preference systems on  $s$ .

## 2.3 Cognitive Partitions

Consider the following definitions:

**Definition 1** *Two mental states  $s, s' \in S$  are similar if they induce the same actions and the same preferences over  $\mathcal{U}_i \times \{0, 1\}$ : That is, if  $a_i^s = a_i^{s'}$  and  $\succsim_s = \succsim_{s'}$ .*

*Two reasoning processes  $\pi, \pi' \in \Pi$  are similar if they induce sequences of pairwise similar states. That is, if  $s^{\pi, k}$  and  $s^{\pi', k}$  are similar for every  $k \in \mathbb{N} \cup \{0\}$ .*

Definition 1 states that two states are similar, if the agent has the same attitude towards the reasoning process as well as the same candidate action in those states. In turn, the next definition states that two games are ‘cognitively similar’ if they induce a sequence of cognitively similar states.

**Definition 2** *Two games  $u, v \in \mathcal{U}$  are cognitively similar (for player  $i$ ), if they induce similar reasoning processes. That is, if  $\pi(u)$  and  $\pi(v)$  are similar.*

Clearly, the relations in definitions 1 and 2 are equivalence relations. We let  $\mathcal{C}$  denote the *cognitive partition*, that is the partition on  $\mathcal{U}$  induced by the cognitive equivalence relation, and let  $C \in \mathcal{C}$  denote a generic cognitive equivalence class. We denote by  $C(u)$  the equivalence class that contains  $u$ . Hence,  $u, v \in \mathcal{U}$  are *cognitively similar* if and only if  $C(v) = C(u)$ .

We impose the following condition on the cognitive partition:

**Condition 2** *For any  $C \in \mathcal{C}$ , there exists  $\alpha \in \Delta(A)$  such that for all  $j \in I$ ,  $\bigcap_{u \in C} BR_j^u(\alpha) \neq \emptyset$ .*

This condition requires a minimal level of strategic similarity between two games to be considered cognitively equivalent. In particular, two games can be cognitively equivalent only if there exists a (possibly correlated) action profile that induces the same best responses for every player in both games.

**Example 1** *Consider the following three games: a ‘Battle of the Sexes’ on the left, a Prisoners’ Dilemma in the middle, and a Matching Pennies game on the right:*

	$L$	$R$		$L$	$R$		$L$	$R$
$U$	3, 1	0, 0	$U$	1, 1	-1, 2	$U$	1, -1	-1, 1
$D$	0, 0	1, 3	$D$	2, -1	0, 0	$D$	-1, 1	1, -1

*Condition 2 allows the first two games to belong to the same equivalence class, but not the third. To see this, notice that profile  $(D, R)$  produces the same best reply for both players in the left and middle game (namely,  $(D, R)$ ), but there is no distribution  $\alpha \in \Delta(A)$  that would generate the same profile of best responses in the game on the right and in one of the others. Essentially, whereas the ‘Battle of the Sexes’ and the Prisoners’ Dilemma are quite different from a strategic viewpoint, they both share a certain element of coordination. The Matching Pennies game is intrinsically a game of conflict, and in this sense it is more different from the other games than they are from each other. Condition 2 captures this minimal notion of strategic similarity within a cognitive equivalence class, allowing (but not requiring) the first two games to be cognitively equivalent, but not the third.*

Notice that, under Condition 2, for every  $C \in \mathcal{C}$  and for every  $i \in N$ , there exists a non-empty set  $M_i(C) \subseteq \Delta(A_{-i})$  of player  $i$ 's conjectures that induce invariant best responses within the cognitive equivalence class:

$$M_i(C) = \left\{ \mu^i \in \Delta(A_{-i}) : \bigcap_{u \in C} BR_i^u(\mu^i) \neq \emptyset \right\}. \quad (1)$$

### 2.3.1 Cognitive Partitions: Discussion and Examples

In the following, we will maintain that the mapping that assigns games to reasoning processes, hence the cognitive partition, is exogenous. The goal of our analysis will be to endogenize the number of steps in this reasoning process performed by the agent, and to identify how the depth of reasoning varies when the payoffs of the game are varied. Here we explain how to think about such exogenous objects of our model.

In an important example of a reasoning process, common in the literature on level- $k$  reasoning, a player's reasoning process is modeled by a sequence of actions  $\{\hat{a}_i^k\}_{k \in \mathbb{N}}$  that correspond to 'increasingly sophisticated' behavior. This will also be our leading example in the applications of our model. For instance, in a symmetric game, the baseline model of Nagel (1995) specifies a level-0 action that represents an exogenous 'anchor' of the reasoning process (it may be a pure or mixed strategy); the level-1 action is the best response to the level-0 action, the level-2 is a best response to level-1, and so forth. If we hypothesize that player  $i$  approaches a specific game  $u$  this way, then the machine  $\pi(u)$  associated to  $u$  generates a sequence of  $\{s^k\}_{k \in \mathbb{N}}$  such that the associated actions  $\{a_i^{s^k}\}_{k \in \mathbb{N}}$  reproduce the sequence entailed by the level- $k$  reasoning: that is,  $a_i^{s^k} = \hat{a}_i^k$  for every  $k$ . In this case, a sensible notion of cognitive partition may be the following:

(C.1) If  $u$  and  $u'$  are such that, for every player, the pure-action best response functions in  $u$  and  $u'$  are the same, then  $C(u) = C(u')$ .

According to (C.1), if  $u'$  has the same pure action best responses as  $u$ , then the reasoning process  $\pi(u')$  generates a sequence of states that produces the same sequence of  $\{a_i^k\}_{k \in \mathbb{N}}$  as game  $u$ . This is a plausible assumption if the sequence  $\{a_i^k\}_{k \in \mathbb{N}}$  is generated by iterating the pure action best responses as in Nagel (1995), and if there are no reasons to believe that the anchor  $a^0$  would be different in  $u'$ . If the anchor changes, then the two games should belong to different cognitive classes, and (C.1) would not be a good assumption.

Other models of level- $k$  reasoning (e.g., Camerer, Ho and Chong, 2004) assume that the level- $k$  action is not a best response to the level- $(k-1)$ , but to a distribution of the lower actions (lower than  $k$ ). In this case, we can think of the level- $k$  path  $\{a_i^k\}_{k \in \mathbb{N}}$  to be generated by a sequence of level- $k$  beliefs  $\mu^k \in \Delta(\{a^0, \dots, a^{k-1}\})$ , such that  $a_i^k \in BR_i^u(\mu^k)$ . A sensible specification of the cognitive partition may then be to require, for two games to be 'cognitively equivalent', that they induce the same best responses to such beliefs  $\{\mu^k\}_{k \in \mathbb{N}}$ . For instance:

(C.2) Given  $\{\mu^k, a_i^k\}_{k \in \mathbb{N}}$  and  $u$  such that  $a_i^k \in BR_i^u(\mu^k)$ , if  $u$  and  $u'$  are such that  $BR_i^u(\mu^k) = BR_i^{u'}(\mu^k)$  for each  $k$ , then  $C(u) = C(u')$ .

Outside of the domain of level- $k$  theories, another form of iterated reasoning is provided by iterated deletion of dominated strategies. These are examples of reasoning processes in which

a large set of best responses are considered at the same time (e.g., the best responses to all distributions concentrated on profiles that have not been deleted yet). In the extreme case in which all the profiles of best responses matter for the reasoning process, then the cognitive partition may be specified in terms of equivalence classes of players' von Neumann-Morgenstern (vNM) preferences. That is, cognitive equivalence classes are invariant to affine transformations of the payoffs of every player.

(C.3) If  $u = (u_j)_{j \in N}$  and  $u' = (u'_j)_{j \in N}$  are such that, for each  $j$ , there exists  $\gamma_j \in \mathbb{R}_+$  and  $m_j \in \mathbb{R}$  such that  $u_j = \gamma_j \cdot u'_j + m_j$ , then  $C(u) = C(u')$ .

To account for the possibility that the reasoning processes may be affected by translations, (C.3) may be weakened by setting  $m_j = 0$  for all  $j$ , so that the cognitive classes would be invariant to linear transformations only, or by further requiring that  $\gamma_j = \gamma$  for all  $j$ , so that the scaling factor is common to all players. These are only some examples of cognitive partitions. As already explained, we maintain the mapping from games to reasoning processes (hence the cognitive partition) as exogenous here. Identifying the 'correct' mapping from games to processes and the corresponding cognitive partition is an important complement to our analysis, which we leave to future research.

## 2.4 Preview of the Cost-Benefit Representation

In the next section we will introduce axioms on the preference systems  $(\succsim_s)_{s \in S}$ , and obtain different representations for our cost-benefit approach. The first six axioms deliver our core result (Theorem 1), which provides foundations to our cost-benefit approach.

**Core Representation:** For any cognitive equivalence class  $C \in \mathcal{C}$ , there exist functions  $W_i : \mathcal{U}_i \times \mathbb{N} \rightarrow \mathbb{R}_+$  and  $c_i : \mathbb{N} \rightarrow \mathbb{R}_+ \cup \{\infty\}$  such that for any game  $u \in C$  and for any  $k \in \mathbb{N}$ ,  $u_i^1 \succsim_{k-1} u_i^0$  if and only if  $W_i(u_i, k) \geq c_i(k)$ . Furthermore, for any  $k \in \mathbb{N}$ , the 'value of reasoning  $W_i(\cdot, k)$  is increasing in the 'payoff differences'  $D(u_i, a_i^{k-1}) := (u_i(a_i, a_{-i}) - u_i(a_i^{k-1}, a_{-i}))_{(a_i, a_{-i}) \in A}$ . That is: for each  $u, u' \in C$ ,  $D(u_i, a_i^{k-1}) \geq D(u'_i, a_i^{k-1})$  implies  $W_i(u_i, k) \geq W_i(u'_i, k)$ , and  $W_i(u_i, k) = 0$  whenever  $D(u_i, a_i^{k-1}) = \mathbf{0}$ .

We refer to this as the 'core representation', which corresponds to the model that we tested experimentally in Alaoui and Penta (2013). Notice that the cost functions are pinned down by the cognitive equivalence classes, and payoff transformations within the same equivalence class only affect the 'incentives to reason' (through the function  $W$ ), but not the 'cost of reasoning'. As previously discussed, the cognitive partition thus plays the role of defining the domain in which exercises of comparative statics are meaningful even in the absence of precise information about the cost function: even if the shape of  $c_i(k)$  is unknown, if  $u, u' \in C$  and  $D(u_i, a_i^{k-1}) \geq D(u'_i, a_i^{k-1})$ , then the agent's cognitive bound is (weakly) higher in game  $u$  than in  $u'$ .

One special case of the 'core representation' that will be used for the analysis of the 'treasures' in Section 5 is obtained by imposing one extra axiom, which requires that the agent is particularly cautious about the validity of his current understanding of the game, hence his disposition to further think about the game is particularly strong. This extra axiom delivers the following functional form for the value of reasoning function:

$$W_i(u_i, k) = \max_{a_{-i} \in A_{-i}} u_i(a_i^*(a_{-i}), a_{-i}) - u_i(a_i^{k-1}, a_{-i}) \geq c(k).$$

In this representation, the value of the next round is equal to the maximum difference between the payoff that the player could get if he chose the optimal action  $a_i^*$  and the payoff he would receive given his current action  $a_i^{k-1}$ , out of all the possible opponent's actions.

### 3 The Value of Reasoning and Cost of Introspection: Axioms and Representations

We impose the following axioms on the preference systems  $(\succsim_s)_{s \in S}$ :

**Axiom 1 (Weak Order)** *Relation  $\succsim_s$  is complete and transitive for each  $s \in S$ .*

The next axiom defines the scope of our theory. It consists of two parts. Part (S.1) states that the reasoning process is purely instrumental in informing the player's choice. Thus, if  $i$ 's payoffs are constant in his own action, the agent would never strictly prefer to think harder. In the representation, this will correspond to a value of reasoning equal to zero. This condition can be equivalently stated as requiring that the agent prefers not to think harder ( $u_i^0 \succsim_s u_i^1$ ) whenever  $u_i$  is such that  $D(u_i) = \mathbf{0}$ , where  $D(u_i)$  denotes the vector of payoff differences in the game, defined as follows: For any  $u_i \in \mathcal{U}_i$ , and  $\hat{a}_i \in A_i$ , let

$$D(u_i, \hat{a}_i) = (u_i(a_i, a_{-i}) - u_i(\hat{a}_i, a_{-i}))_{(a_i, a_{-i}) \in A},$$

and  $D(u_i) = D(u_i, \hat{a}_i)_{\hat{a}_i \in A_i}.$

Hence, 'positive payoff differences' are a necessary condition to induce the agent to reason. Part (S.2) pushes the idea further, requiring that the incentives to reason are completely driven by such payoff differences: games with the same payoff differences provide the same incentives to reason.

**Axiom 2 (Scope)** *For each  $s \in S$ :*

*S.1 [Instrumentality] If  $u_i$  is constant in  $a_i$ , then  $u_i^0 \succsim_s u_i^1$ . That is,  $D(u_i) = \mathbf{0}$  implies  $u_i^0 \succsim_s u_i^1$ .*

*S.2 [Incrementality] If  $D(u_i) = D(v_i)$ , then  $u_i^1 \succsim_s u_i^0$  if and only if  $v_i^1 \succsim_s v_i^0$ .*

For convenience, we let  $\hat{U}_i$  denote the set of payoff functions that are constant in own action:  $\hat{U}_i = \{u_i \in \mathcal{U}_i : u_i \text{ is constant in } a_i\}$ .

The next axiom is a 'cost-independence' condition, requiring that the extra incentives that are required to reason about a constant game in a given mental state are independent of the game itself.

**Axiom 3 (Cost-Independence)** *For each  $s \in S$ , for any  $u_i, v_i \in \hat{U}_i$  and  $c \in \mathbb{R}_+$ ,  $(u_i + c, 1) \sim_s (u_i, 0)$  if and only if  $(v_i + c, 1) \sim_s (v_i, 0)$ .*

To understand the previous axiom, notice that (S.1) implies that  $u_i^1 \succsim_s u_i^0$  whenever  $u_i \in \hat{U}_i$ . We may thus have two cases: If  $u_i^1 \sim_s u_i^0$ , that is if the agent, at mental state  $s$ , is indifferent between reasoning and not reasoning further about a constant game, then the axiom requires that he would also be indifferent between reasoning and not reasoning in any other constant game; If instead the agent strictly prefers not to reason about a constant game  $u_i \in \hat{U}_i$ , but he could be

made indifferent if reasoning was accompanied by an extra reward  $c$ , then that reward would also make him indifferent between reasoning or not in any other constant game. Hence, the reward required to incentivize the agent to reason about a constant game (which, by axiom S.1, provides no intrinsic incentives to reason) depends only on the mental state, not on the payoffs. Axiom 3 will be used as a ‘calibration axiom’ to pin down the cost of reasoning associated to a given mental state.

The following axiom contains standard properties:

**Axiom 4** For each  $s \in S$ :

*M* [ $\succsim_s^1$ -**Monotonicity**]  $u_i \geq v_i$  implies  $u_i^1 \succsim_s v_i^1$  and there exists  $a'_i$  s.t. if  $u_i \geq v_i$  and  $u_i(a'_i, a_{-i}) \geq v_i(a'_i, a_{-i})$ , then  $u_i^1 \succ_s v_i^1$ .

*C.1* [ $\succsim_s^1$ -**Archimedean**] For any  $v_i, u_i, h_i \in \mathcal{U}_i$  such that  $u_i \succ_s^1 v_i \succ_s^1 h_i$ , there exist  $0 \leq \beta \leq \alpha \leq 1$  such that  $\alpha u_i + (1 - \alpha) h_i \succ_s^1 v_i \succ_s^1 \beta u_i + (1 - \beta) h_i$ .

*C.2* For any  $u_i \in \mathcal{U}_i$ , if there exists  $c \in \mathbb{R}_+$  s.t.  $(u_i + c, 1) \succ_s (u_i, 0) \succ_s (u_i, 1)$ , then there exists  $c^* \in [0, c)$  such that  $(u_i, 0) \sim_s (u_i + c^*, 1)$ .

Parts M and C.1 are standard Monotonicity and Archimedean axioms for the  $\succsim_s^1$  preference system over  $\mathcal{U}_i \times \{1\}$ .<sup>6</sup> Part C.2 instead is a minimal continuity requirement for the comparison of games with and without further reasoning. The axiom is particularly weak. For instance, it allows non-continuous preferences such as  $(u_i, 0) \succ_s (v_i, 1)$  for all  $u_i, v_i \in \mathcal{U}_i$ . That is, an agent who cannot be incentivized to perform the next step of reasoning, no matter how the payoffs are changed. This is a lexicographic ordering, which can be accommodated by our framework, and would correspond to the case of an infinite cost of reasoning at that mental state.

For any  $u_i \in \mathcal{U}_i$  define the set

$$\Lambda(u_i) = \left\{ v_i \in \mathcal{U}_i : \begin{array}{l} \exists \lambda : A_i \rightarrow A_i \text{ such that} \\ u_i(\lambda(a_i), a_{-i}) = v_i(a_i, a_{-i}) \text{ for all } a \in A \end{array} \right\}. \quad (2)$$

In words, the set  $\Lambda(u_i)$  comprises all the games that can be obtained from  $u_i$  by relabeling acts, or (if the relabeling function  $\lambda$  is not onto) by dropping some acts from the game.

The next axiom concerns the preference system  $\succsim_s^1$ , and requires that if the agent were to play a game after having performed the next step of reasoning, then the agent weakly prefers to play the game about which he has been reasoning, rather than a ‘relabeling’ of that game. The axiom can be seen as requiring that the reasoning process is apposite to the specific game, and this is accounted for in the way the agent forms his outlook about the extra steps of reasoning at a mental state.

**Axiom 5 (1-Unimprovability)** For any  $s \in S$ , for any  $u$  such that  $s \in S(\pi(u))$  and for any  $v_i \in \Lambda(u_i)$ ,  $u_i \succsim_s^1 v_i$ .

We view axioms 1-5 as the basic axioms of our framework, those that define the scope of the theory and the determinants of the value and cost of reasoning. These axioms are silent about

<sup>6</sup>As usual, the Archimedean property is required for the richness of the choice space ( $\mathcal{U}_i$  is uncountable). It could be weakened to a separability axiom for the general representation (Theorem 1). We maintain it nonetheless because it will be required for the expected utility extension that we introduce in the next subsection.

the content of the reasoning process itself. The next axiom essentially states that the output of the reasoning process, as evaluated by the preference system  $\succsim_s^0$ , consists of the action  $a_i^s$  that the agent views as the most sophisticated at state  $s$ .

**Axiom 6 (Reasoning output)** *For each  $s \in S$ ,  $u_i \sim_s^0 u_i^s$  whenever  $u_i^s$  is such that  $u_i^s(a_i, a_{-i}) = u_i(a_i^s, a_{-i})$  for all  $(a_i, a_{-i}) \in A$ .*

This is consistent with the idea of level- $k$  models that the output of each step of reasoning is a specific action. The framework defined by axioms 1-5, however, can accommodate other forms of reasoning as well. For instance, rationalizability or DK-models of iterated dominance, can be thought of as reasoning processes in which the output consists of both an action and a set (e.g., the support of the feasible conjectures). Such alternative forms of reasoning could be accommodated through suitable modifications of Axiom 6.

The next theorem provides a representation theorem for the individuals' choices to continue reasoning about the game. It states that each equivalence class  $C$  determines a cost function  $c : \mathbb{N} \rightarrow \mathbb{R}_+$ . For each  $k \in \mathbb{N}$ ,  $c(k)$  represents the 'cognitive cost' of performing the  $k$ -th step of reasoning (given the previous steps). This cost is weighed against the value  $W(u_i, k)$  of performing the  $k$ -th step of reasoning in the game  $u \in C$  in the cognitive equivalence class. Notice that, by the definition of the cognitive partition (Def. 2), once  $C$  is fixed, any  $u, v \in C$  induce sequences  $\{s^k(\pi(u))\}_{k \in \mathbb{N}}$  and  $\{s^k(\pi(v))\}_{k \in \mathbb{N}}$  that correspond to the same action  $a_i^{s^k(\pi(u))}$  and preferences  $\succsim_{s^k(\pi(u))}$  for each  $k$ . Once the cognitive class is clear, we can thus simplify the notation and refer to such sequences as  $a_i^k$  and  $\succsim_k$ .

**Theorem 1** *Under Axioms 1-6, for any cognitive equivalence class  $C \in \mathcal{C}$ , there exist functions  $W : \mathcal{U}_i \times \mathbb{N} \rightarrow \mathbb{R}_+ \cup \{\infty\}$  and  $c : \mathbb{N} \rightarrow \mathbb{R}_+ \cup \{\infty\}$  such that for any game  $u \in C$  and for any  $k \in \mathbb{N}$ ,  $u_i^1 \succ_{k-1} u_i^0$  if and only if  $W(u_i, k) \geq c(k)$ . Furthermore, for any  $k \in \mathbb{N}$ : for each  $u, u' \in C$ ,  $D(u_i, a_i^{k-1}) \geq D(u'_i, a_i^{k-1})$  implies  $W(u_i, k) \geq W(u'_i, k)$ , and  $W(u_i, k) = 0$  whenever  $D(u_i, a_i^{k-1}) = \mathbf{0}$ .*

### 3.1 Further Restrictions

We present next two further representation theorems, obtained imposing additional restrictions besides those entailed by axioms 1-6. These representations will impose more structure on the 'value of reasoning' function  $W$ . Both representations will suggest an intuitive interpretation of the determinants of the value of reasoning, in terms of beliefs about the opportunity to improve upon the current understanding of the game through introspection. The first representation is obtained adding a standard independence axiom for the  $\succsim_s^1$  preference systems, and delivers a representation in which the agent's attitude towards the value of reasoning resembles a standard 'value of information', except that individuals have subjective beliefs about the realization of 'signals'. The second representation instead follows from an axiom that captures the attitude towards reasoning of an agent who is always cautious on relying on his current understanding of the game, and therefore has a maximal disposition to think more.

#### 3.1.1 EU-attitude towards reasoning

We introduce next a standard independence axiom for the  $\succsim_s^1$  preferences:

**Axiom 7 (1-Independence)** For each  $s \in S$ : For all  $v_i, u_i, h_i \in \mathcal{U}_i$ ,  $u_i \succ_s^1 v_i$  if and only if  $\alpha u_i + (1 - \alpha) h_i \succ_s^1 \alpha v_i + (1 - \alpha) h_i$  for all  $\alpha \in [0, 1)$ .

This axiom plays the usual role of inducing a property of linearity (in this case of the ‘value of reasoning’ function) which allows for a representation that can be interpreted in terms of ‘beliefs’ and ‘expected value of reasoning’.

**Theorem 2** Under Axioms 1-6 and 7, for any cognitive equivalence class  $C \in \mathcal{C}$ , there exists  $c : \mathbb{N} \rightarrow \mathbb{R}_+ \cup \{\infty\}$  and for any  $k \in \mathbb{N}$  a simple probability distribution  $p^k \in \Delta(M_i(C))$  such that for any game  $u \in C$ ,  $u_i^1 \succ_{k-1} u_i^0$  if and only if

$$\sum_{\mu \in \Delta(A_{-i})} p^k(\mu) \sum_{a_{-i}} \mu(a_{-i}) [u_i(a_i^*(\mu), a_{-i}) - u_i(a_i^{k-1}, a_{-i})] \geq c(k), \quad (3)$$

where  $a_i^*(\mu) \in BR_i^u(\mu)$  for any  $\mu \in M_i(C)$ .

In this representation, the ‘value of reasoning’ has an intuitive interpretation. It is *as if* the agent has beliefs  $p^k \in \Delta(M_i(C))$  about the outcome of his future step of reasoning. Such outcomes consist of conjectures about the opponent’s behavior, to which the agent will be able to respond optimally, improving on his previous understanding of the game (as entailed by the ‘current action’  $a_i^{k-1}$ ). The left-hand side of equation 3 is reminiscent of the standard notion of ‘expected value of information’ often used in information economics. However, it should be noticed that the beliefs in the left-hand side of equation (3) are derived from the  $\succ_s^1$ -preference system used to describe the agent attitudes towards the future steps of reasoning, which is unrelated to the determinants of the current understanding of the game  $a_i^{k-1}$ . In contrast, standard models of information economics require that agents are ‘Bayesian’ in the stronger sense of using a single prior over everything that affects choices. As such, the beliefs that describe the outlook on future steps of reasoning should be consistent with current choice  $a_i^{k-1}$ . Formally, let  $\hat{\mu}^k \in \Delta(A_{-i})$  be defined from the representation in Theorem 2 so that  $\hat{\mu}^k(a_{-i}) = \sum_{\mu \in \Delta(A_{-i})} p^k(\mu) \mu(a_{-i})$  for each  $a_{-i} \in A_{-i}$ . Then  $\hat{\mu}^k \in BR^{-1}(a_i^{k-1})$  for each  $u_i \in C$ . This property can be obtained appending Axioms 1-7 with the following, which requires that the current understanding of the game cannot be improved upon from the viewpoint of the preferences describing the outlook on the future steps of reasoning:

**Axiom 8 (1-0 Consistency)** For each  $u \in \mathcal{U}$  and for each  $s \in S(\pi(u))$ ,  $u_i^s \succ_s^1 v_i^s$  for all  $v_i \in \Lambda(u_i)$ .

This property, natural in standard information problems, is not necessarily desirable in the present context. Axiom 8 requires that the agent’s attitude towards the future steps of reasoning is essentially the same as that entailed by his current understanding of the game, which seems too narrow to accommodate central features of boundedly rational reasoning. In particular, it fails to capture the idea that the agent may not be fully aware of the determinants of the reasoning process he has yet to perform. Such ‘unawareness’ would introduce a disconnect between the heuristics that describe the agent’s attitude towards future steps of reasoning and those that the agent has already performed. This feature would be inconsistent with Axiom 8. For this reason, we do not include Axiom 8 among the axioms of the EU-representation. For the sake of completeness, we nonetheless provide the following result:

**Proposition 1** Under Axioms 1-7 and 8, the representation in Theorem 2 is such that, for each  $C \in \mathcal{C}$ , for each  $u \in C$  and for each  $k$ ,  $a_i^{k-1} \in BR^u(\hat{\mu}^k)$ .

### 3.1.2 Optimism towards reasoning

One particular way in which bounded rationality may determine a disconnect between current understanding and evaluating future steps of reasoning is that the agent may be particularly cautious or pessimistic about the validity of his current understanding of the game. Consequently, his disposition to further think about the game is particularly strong. The following representation is obtained appending an axiom that captures precisely this idea of ‘optimism towards reasoning’ (or ‘pessimism towards current understanding’), in the sense that the agent’s outlook towards the future reasoning process raises a strong challenge to his current understanding. From this viewpoint, the following representation stands in sharp contrast to the idea implicit in Axiom 8, which can be interpreted as saying that the agent’s outlook towards the future does not challenge the optimality of the current understanding of the game.

Before introducing the axiom that captures the insight we have just discussed, it is useful to consider the following result:

**Lemma 1** *Under Axioms 1-6,  $u_i - u_i^s \succsim_s^1 v_i - v_i^s$  if and only if  $v_i^1 \succsim_s v_i^0$  implies  $u_i^1 \succsim_s u_i^0$ .*

According to Lemma 1, in every mental state, it is possible to define a binary relation that orders payoff functions in terms of the strength of the incentives to reason that they provide. We define such binary relation as follows:  $u_i \succsim_s^* v_i$  if and only if  $u_i - u_i^s \succsim_s^1 v_i - v_i^s$ . Notation  $u_i \succsim_s^* v_i$  means ‘in state  $s$ , payoffs  $u_i$  provide stronger incentives to reason than  $v_i$ ’.

For any  $u_i \in \mathcal{U}_i$ , let :

$$\Gamma(u_i) = \left\{ v_i \in \mathcal{U}_i : \begin{array}{l} \exists \gamma : A \rightarrow A \text{ such that} \\ u_i(\gamma(a)) = v_i(a) \text{ for all } a \in A \end{array} \right\}. \quad (4)$$

The set  $\Gamma(u_i)$  is analogous to the set  $\Lambda(u_i)$  introduced in eq. (2), except that payoffs in  $\Gamma(u_i)$  are obtained through a relabeling of the entire action profiles, not only of player  $i$ ’s own action.

The next axiom states that, whenever  $s \in S(\pi(u))$ , then the agent has stronger incentives to reason if payoffs are as in the ‘actual game’  $u_i$  than he would if payoffs were modified through a relabeling of the strategy profiles.

**Axiom 9** *For and  $s \in S$  and for any  $u \in \mathcal{U}$  such that  $s \in S(\pi(u))$ ,  $u_i \succsim_s^* v_i$  for all  $v_i \in \Gamma(u_i)$ .*

**Theorem 3** *Under Axioms 1-6 and 9, for any cognitive equivalence class  $C \in \mathcal{C}$ , there exist a function  $c : \mathbb{N} \rightarrow \mathbb{R}_+ \cup \{\infty\}$  such that for any game  $u \in C$  and for any  $k \in \mathbb{N}$ ,  $u_i^1 \succsim_{k-1} u_i^0$  if and only if*

$$\max_{a_{-i} \in A_{-i}} u_i(a_i^*(a_{-i}), a_{-i}) - u_i(a_i^{k-1}, a_{-i}) \geq c(k).$$

## 4 Endogenous Level-k Reasoning

In this section we introduce a special case of the model of reasoning developed above, which extends existing models of level-k reasoning by endogenizing individuals’ depth of reasoning. Consistent with the representation theorems above, players’ depth of reasoning stems from a cost-benefit analysis, in which individuals’ cognitive abilities interact with the incentives provided by the game payoffs. Furthermore, players take into consideration that their opponents follow a similar process of reasoning when choosing an action. Thus, in our model, the rounds of introspection that a

player performs can be disentangled from the rounds he believes his opponents perform. Since a player can face opponents of different skills, he may in principle conceive of an opponent who is as sophisticated, or more, than he is himself. Our model permits this eventuality, and resolves this apparent conceptual difficulty of the level- $k$  approach. Hence, by making explicit an appealing feature of level- $k$  models that play follows from a reasoning procedure, our framework may be instrumental in attaining a deeper understanding of the mechanisms behind this approach, and paves the way to a more complete model of procedural rationality.

In the Section 5 we further show that, combined with the restrictions on the value of reasoning derived in the previous section, this model provides a unified framework to explain important experimental evidence in games of initial responses, such as the famous ‘little treasures’ of Goeree and Holt (2001).

#### 4.1 Individual Reasoning

To keep the notation simple, we focus on two players games,  $G = (A_i, u_i)_{i=1,2}$ . As in the previous section, we analyze each player in isolation. The ‘*path of reasoning*’ of player  $i$  is described by a sequence of (possibly mixed) strategy profiles  $\{a^k\}_{k \in \mathbb{N} \cup \{0\}}$ . Profile  $a^0 = (a_1^0, a_2^0)$  is referred to as ‘the anchor’ and represents the way player  $i$  approaches the game. We follow the literature on level- $k$  reasoning in assuming that the path of reasoning is determined by iterating players’ best responses. That is, action  $a_i^0$  represents  $i$ ’s default action, if he doesn’t think at all. If  $i$  performs one step of reasoning, then he conceives that  $j$ ’s default action would be  $a_j^0$ , hence he best responds with  $a_i^1 = BR_i(a_j^0)$ , and so on. In general, the best response need not be unique. In case of multiplicity, we assume that the action is drawn from a uniform distribution over the best responses. We abuse notation and write  $a_i^1 = BR_i(a_j^0)$  in both cases. We thus define recursively:

$$a_i^k = BR_i(a_j^{k-1}) \tag{5}$$

$$\text{and } a_j^k = BR_j(a_i^{k-1}). \tag{6}$$

In general, we may have two cases: If the anchor  $a^0$  is a Nash Equilibrium, equations (5-6) imply that  $a^k = a^0$  for every  $k$ . That is, if player  $i$  approaches a game with an ‘anchor’ that specifies a certain equilibrium, further introspection of that initial presumption would not challenge his initial view. If instead  $a^0$  is not a Nash Equilibrium, then  $a_i^k \neq a_i^{k+1}$  and the reasoning process generates a path that may converge to a Nash equilibrium or enters a loop that would only be interrupted by the ‘choice’ of stopping the reasoning once it becomes too costly (or by the agent’s beliefs about the opponent’s choice):

**Example 2** Consider the following game, parameterized by  $x \in \mathbb{R}$ :

	$L$	$R$
$T$	$x, 40$	$40, 80$
$B$	$40, 80$	$80, 40$

If  $x > 40$ , this is a version of the matching pennies game. If player 1 approaches the game with anchor  $a^0 = (T, R)$ , then his ‘*path of reasoning*’ is  $\{a_1^k\}_{k \in \mathbb{N} \cup \{0\}} = (T, T, B, B, T, T, \dots)$  (it follows from equations 5 and 6) If the anchor instead is  $a^0 = (T, L)$ , then the path of reasoning would be  $\{a_1^k\}_{k \in \mathbb{N} \cup \{0\}} = (T, B, B, T, T, BB, \dots)$ , and so on. These sequences clearly do not converge. If

$x < 40$ , then action  $T$  is dominated or player 1, and the only rationalizable strategy profile (hence, the only Nash equilibrium) in this game is  $(B, L)$ . If player 1 approaches the game with  $(B, L)$  as an anchor, then the path of reasoning is  $\{a_1^k\}_{k \in \mathbb{N} \cup \{0\}} = (B, B, \dots)$ . If, for instance, player 1 approaches the game with  $(T, L)$  as an anchor, then his path of reasoning is  $\{a_1^k\}_{k \in \mathbb{N} \cup \{0\}} = (T, T, B, B, B, B, \dots)$ . In this case, the path of reasoning converges to an equilibrium even if the anchor is not one.

Consistent with the model of the reasoning process developed in Sections 2 and 3, we assume that player  $i$ 's understanding of the game is determined by a cost-benefit analysis, where the cost and the value of reasoning are represented by functions  $c_i : \mathbb{N} \rightarrow \mathbb{R}$  and  $W_i : \mathbb{N} \rightarrow \mathbb{R}$ , respectively. Consistent with Theorem 1, player  $i$  therefore stops the reasoning process when the value of performing an additional round of introspection exceeds the cost. The point at which this occurs identifies his cognitive bound  $\hat{k}_i$ .

It is useful to introduce the following mapping, which identifies the intersection between the value of reasoning and the cost function: Let  $\mathcal{K} : \mathbb{R}_+^{\mathbb{N}} \times \mathbb{R}_+^{\mathbb{N}} \rightarrow \mathbb{N}$  be such that, for any  $(c, W) \in \mathbb{R}_+^{\mathbb{N}} \times \mathbb{R}_+^{\mathbb{N}}$ ,

$$\mathcal{K}(c, W) = \min \{k \in \mathbb{N} : c(k) \leq W(k) \text{ and } c(k+1) > W(k+1)\}, \quad (7)$$

with the understanding that  $\mathcal{K}(c, W) = \infty$  if the set in equation (7) is empty. Player  $i$ 's *cognitive bound*, which represents his understanding of the game, is then determined by the value that this function takes at  $(c_i, W_i)$ :

**Definition 3** Given cost and value functions  $(c_i, W_i)$ , the cognitive bound of player  $i$  is defined as:

$$\hat{k}_i = \mathcal{K}(c_i, W_i). \quad (8)$$

Notice that the cognitive bound  $\hat{k}_i$  is monotonic in the level of the cost and of the value functions, irrespective of the shape of these functions:  $\hat{k}_i$  (weakly) decreases as the cognitive costs increase, and it (weakly) increases as the value of reasoning increases.

**Remark 1** Under the maintained assumptions of Theorem 1:

1. For any  $c_i$ , if  $W_i'(k) \geq W_i(k)$  for all  $k$ , then  $\mathcal{K}(c_i, W_i') \geq \mathcal{K}(c_i, W_i)$ .
2. For any  $W_i$ , if  $c_i'(k) \geq c_i(k)$  for all  $k$ , then  $\mathcal{K}(c_i', W_i) \leq \mathcal{K}(c_i, W_i)$ .

**Example 3** In the representation of Theorem 3, for instance, the value of reasoning takes the following shape:

$$W_i(k) = \max_{(a_i, a_{-i}) \in A} u_i(a_i^*(a_{-i}), a_{-i}) - u_i(a_i^{k-1}, a_{-i}).$$

In the game of example 2, with  $x = 80$ , this representation delivers a constant value of reasoning equal to 40: as we have seen, from  $k = 2$  on the path of reasoning of player 1 determines a cycle between  $T$  and  $B$ . Independent of whether  $a_1^{k-1} = T, B$ , however, this representation yields a value  $W_1(k) = 40$  when parameter  $x = 80$ . Clearly, if all payoffs were multiplied by a constant  $X > 1$ , then the value of reasoning would increase to  $W_1'(k) = 40X$ . Figure 1 illustrates how such an increase of payoffs may increase the depth of reasoning of player 1.<sup>7</sup>

<sup>7</sup>Given the definition of  $\mathcal{K}(\cdot)$ , it is easy to verify that the cognitive bound of player 1 would (weakly) increase when  $W$  increases independent of the shape of the cost of reasoning. In particular,  $c_1$  need not be monotonic.

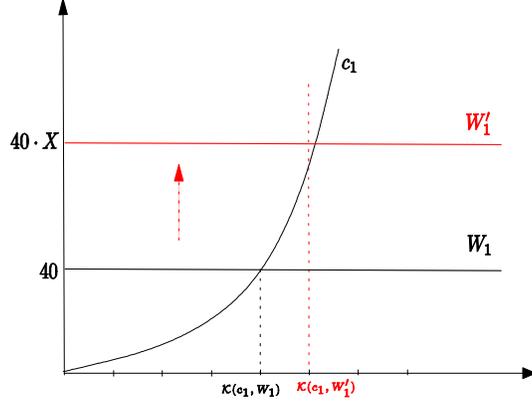


Figure 1: Individual Reasoning

## 4.2 From Reasoning To Choice

To determine player  $i$ 's behavior, given his understanding of the game, we need to relate players' reasoning process to their beliefs about the opponent's reasoning process. We assume that players reason about the game taking into account that their opponent follows a similar cost-benefit procedure. Players therefore take into account the sophistication of their opponents. Since a player's reasoning ability in this model is captured by the cost function, we use the same tool to model players' beliefs over others' sophistication. We specifically define sophistication in the following manner:

**Definition 4** Consider two cost functions,  $c'$  and  $c''$ . We say that cost function  $c'$  corresponds to a 'more sophisticated' player than  $c''$ , if  $c'(k) \leq c''(k)$  for every  $k$ .

For any  $c_i \in \mathbb{R}_+^{\mathbb{N}}$ , it is useful to introduce the following notation:

$$C^+(c_i) = \{c \in \mathbb{R}_+^{\mathbb{N}} : c_i(k) \geq c(k) \text{ for every } k\} \text{ and} \quad (9)$$

$$C^-(c_i) = \{c \in \mathbb{R}_+^{\mathbb{N}} : c_i(k) \leq c(k) \text{ for every } k\}. \quad (10)$$

Sets  $C^+(c_i)$  and  $C^-(c_i)$  respectively comprise the cost functions that are more and less sophisticated than  $c_i$ . Given  $c_i, c'_i \in C^+(c_i)$  and value of reasoning  $W_i$ , we say that ' $c'_i$  is strictly more sophisticated than  $c_i$ , given  $W_i$ ', if  $\mathcal{K}(c'_i, W_i) > \mathcal{K}(c_i, W_i)$ .

We represent players' beliefs about the sophistication of the opponents (as well as their beliefs about the opponent's beliefs) in the standard way, by means of type spaces. The only difference is that each type will be associated to both beliefs and a cost function:

**Definition 5** A  $C$ -type space (or just type space) is defined by a tuple  $\langle (T_j, \gamma_j, \tau_j)_{j \in N} \rangle$  such that  $T_j = \{t_j^1, \dots, t_j^{m_j}\}$  is the set of types of player  $l$ ,  $\gamma_j : T_j \rightarrow \mathbb{R}_+^{\mathbb{N}}$  assigns to each type a cost function and  $\tau_j : T_j \rightarrow \Delta(T_j)$  assigns to each type a belief about the opponent types. (For convenience, we write  $\tau_{t_j}$  instead of  $\tau_j(t_j)$ .)

The simplest case of cognitive type space is the 'independent common prior' CTS, in which  $\gamma_l$ 's are one-to-one and for each  $l$  there exist priors  $p_l \in \Delta(T_l)$  and for all  $t_l, t'_l \in T_l$ ,  $\tau_{t_l} = \tau_{t'_l} = p_{-l}$ . In

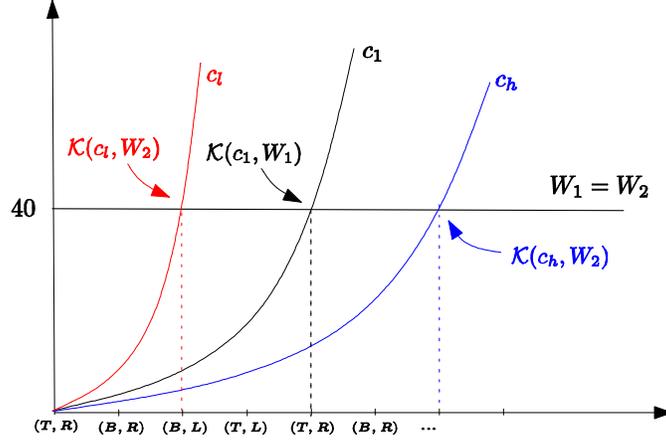


Figure 2: From Reasoning to Choice: the path of reasoning is generated by anchor  $a^0 = (T, R)$ ; choice depends on the beliefs about the opponent's type.

this case, the type space is merely a collection  $T_l = \{c_{t_1}, \dots, c_{t_{m_i}}\}$  of cost functions for each player  $l$ , drawn from some (commonly known) prior distribution  $p_l \in \Delta(T_l)$ .

Player  $i$ 's choice in the game stems from his depth of reasoning, his anchor  $a^0$ , and his beliefs represented by his type  $t_i$  in a type space. The unit of our analysis will thus be a pair  $(t_i, a^0) \in T_i \times A$ .

**Definition 6** Given a type space, and a pair  $(t_i, a^0) \in T_i \times A$ . For each  $j$  define functions  $\alpha_j : T_j \rightarrow A_j$  such that  $\alpha_j^0(t_j) = a_j^0$  for each  $t_j \in T_j$ . Recursively, define for each  $j = 1, 2$ , for each  $t_j \in T_j$  and for each  $k \in \mathbb{N}$ :

$$\alpha_j^k(t_j) = \begin{cases} \alpha_j^{k-1}(t_j) & \text{if } k > \mathcal{K}(c_{t_j}, W_j) \\ BR_j \left( \sum_{t_{-j} \in T_{-j}} \tau_{t_j}(t_{-j}) \cdot \alpha_{-j}^{k-1}(t_{-j}) \right) & \text{otherwise} \end{cases} \quad (11)$$

The choice of type  $t_i$ , given anchor  $a^0$ , is  $\hat{a}_i(t_i, a^0) := \alpha_i^{\mathcal{K}(c_i, W_i)}(t_i)$ .

**Example 4** Consider a type space in which  $T_1 = \{c_1\}$  and  $T_2 = \{c_l, c_h\}$ , such that  $c_l \in C^-(c_1)$ ,  $c_h \in C^+(c_1)$  and the type of player 2 is  $l$  with probability  $q$ . That is, player 1 believes that with probability  $q$  the opponent would be less sophisticated than himself, and with probability  $(1 - q)$  he will be more sophisticated. Consider the game in Example 2 again, with  $x = 80$  so that the value of reasoning is constant and equal to 40 for both players (see Example 3). Figure 2 represents this situation for a player that approaches the game with anchor  $a^0 = (T, R)$ . Given the cost functions and the incentives to reason, we have that  $\mathcal{K}(c_l, W_2) = 2$ ,  $\mathcal{K}(c_1, W_1) = 4$  and  $\mathcal{K}(c_h, W_2) = 6$ . Hence, the depth of reasoning of player 1 is four, which corresponds to action  $a_1^4 = T$  in his path of reasoning. This, however, need not be the action that he plays in general. To see this, notice that the recursion in (11) implies that  $\alpha_2^k(l) = R$  for all  $k \geq 2$ . Hence, if  $q_l = 1$  (that is, if player 1 is certain to play against the less sophisticated opponent), for all  $k \geq 3$   $\alpha_1^k(t_1)$  is equal to the best response to  $R$ , that is  $\alpha_1^k(t_1) = B$  for all  $k \geq 3$ , hence player 1 plays like a level-3,  $B$ . Given the game payoffs, it can be verified that this would be the case as long as  $q_l > 1/2$ . Hence, if player 1 attaches at least probability  $1/2$  to his opponent's depth of reasoning being two, then even though

1's depth of reasoning is four, he would play like a level-3 player. If instead  $q_l < 1/2$ , then the path  $\alpha_1^k(t_1)$  defined by equation (11) coincides with the 'path of reasoning' for all  $k \leq \mathcal{K}(c_1, W_1)$ . Hence, if player 1 believes that his opponent is more sophisticated, then he will play according to his own cognitive bound. The same figure is useful to analyze how player 2 would play if he approached the same game with anchor  $a^0 = (T, R)$ . The low type of player 2 would play according to his own cognitive bound, hence he would play  $R$ . Now let's consider the high type of player 2, given anchor  $a^0 = (T, R)$ . He anticipates that player 1 anticipates that the low type would play  $R$ . If  $q_l > 1/2$ , equation (11) implies that  $\alpha_1^k(t_1) = B$  for all  $k \geq 3$ , that is, player 1 behaves as a level-3. Given this, (11) further implies that  $\alpha_2^k(h) = R$  for all  $k \geq 4$ . Hence the high type behaves as a level-4. If instead  $q_l < 1/2$ , then the high type anticipates that player 1 would play according to his cognitive bound, that is four. Equation (11) then implies that  $\alpha_2^k(h) = L$  for all  $k \geq 5$ . In that case, the high type behaves as a level-5.

## 5 Five 'Little Treasures' of Game Theory

In an influential paper, Goeree and Holt (2001, henceforth GH) conduct a series of experiments on initial responses in different games. While the games they consider are quite different from one another, the experimental design is similar across the games. For each of these games, GH contrast individuals' behavior in the baseline game, or 'treasure', with that observed in a similar game, obtained through simple changes of the payoffs. The latter treatments are referred to as 'contradictions'. For each of these games, GH show that classical equilibrium predictions often perform well in the treasure but not in the contradiction. While most of the contradictions seem intuitive, no unified explanation of these findings has been provided thus far. Among the more surprising aspects of our theory has been its aptitude to provide such a unifying framework for all the games in GH that fall within its domain, namely, the static games with complete information. We next review GH's findings for these games. We then show that the model is not only qualitatively in line with these treasures, but that in fact it is consistent with the empirical distributions of the findings.

### 5.1 Five 'Little Treasures': Review

**Traveler's Dilemma.** In this version of Basu's well-known Traveler's dilemma game, two players choose a number between 180 and 300 (inclusive). The reward they receive is  $\min\{a_1, a_2\} + R$  for the player who gives the lower number, and  $\min\{a_1, a_2\} - R$  for the player who gives the higher number. As long as  $R > 0$ , this game is dominance solvable, and 180 is the only equilibrium strategy. GH consider two cases,  $R = 5$  and  $R = 180$ .<sup>8</sup> They observe that for  $R = 5$ , roughly 80% of subjects choose the highest claim, while for  $R = 180$ , roughly 80% choose the Nash strategy.

**Matching Pennies.** GH consider the following version of the Matching Pennies game. In the 'treasure' treatment, parameter  $x$  is set equal to 80. The game therefore is symmetric, and players uniformly randomize over their actions in the unique Nash equilibrium of the game.

---

<sup>8</sup>GH do not specify the rule in case of tie. In the following, we assume that there are no transfers in that case.

$x = 80$	L (48)	R (52)
T (48)	$x, 40$	40,80
B (52)	40,80	80,40

The empirical distributions observed for the ‘treasure’ treatment (shown in parenthesis) are remarkably close to the equilibrium predictions.

In the ‘contradiction’ treatment, GH modify the parameter  $x$ , setting it equal to 320 in one case and 44 in another. From a Nash equilibrium viewpoint, an agent’s mixture in the equilibrium in this game is chosen so as to make the other indifferent. So, since  $x$  does not affect the payoffs of the column player, the equilibrium distribution of the row player in the contradictions should still be the uniform. Equilibrium therefore entails that it should be the column player that reacts to changes in the row player’s payoffs, reducing the weight on  $L$  when  $x$  increases, and vice versa. This prediction seems counterintuitive, and is indeed rejected by the data: in both cases, more than 90% of the row player choose the action with the relatively higher payoffs,  $T$  when  $x = 320$  and  $B$  when  $x = 44$ . Moreover, this behavior seems to have been anticipated by the column players, with roughly 80 percent of subjects playing the best response to the action played by most of the row players, which is  $R$  when  $x = 320$  and  $L$  when  $x = 44$ .

$x = 320$	L (16)	R (84)	$x = 44$	L (80)	R (20)
T (96)	$x, 40$	40,80	T (8)	$x, 40$	40,80
B (4)	40,80	80,40	B (92)	40,80	80,40

**Coordination Game with a Secure Outside Option.** The next game is a coordination game with one efficient and one inefficient equilibrium, which pay (180, 180) and (90, 90), respectively. The column player, however, has a secure option which pays 40 independent of the row player’s choice.

	L	H	S
L	90,90	(0,0)	$x, 40$
H	0,0	(180,180)	0,40

The parameter  $x$  affects the row player’s payoffs, but not the column player. In the treasure treatment,  $x$  is set equal to 0, so that the secure option of the column player leaves the row player indifferent between his two alternatives. In the contradiction instead the parameter  $x$  is set equal to 400. Notice that, independent of  $x$ , a 50-50 combination of  $H$  and  $L$  dominates  $S$  for the column player. A rational column player therefore would never play  $S$ , and therefore a change in  $S$  should not affect the behavior of the row player, if the latter thinks that the column player is rational. Having eliminated  $S$ , the remaining game has three Nash equilibria:  $(L, L)$ ,  $(H, H)$ , and one in which player randomize putting  $2/3$  probability on  $L$ . So, the change in  $x$  has no impact on the equilibria of the set. However, the magnitude of  $x$  may affect the coordination process. GH do not report the entire distributions for this experiment. The next two matrices summarize the

experimental data as reported by GH:<sup>9</sup>

$x = 0$	L	H (84)	S	$x = 400$	L	H	S
L				L	(16)		
H (96)		(80)		H		(32)	

(12)

**A Minimum-Effort Coordination Game.** In this game players choose efforts levels  $a_1, a_2$  which can be any integers between 110 and 170. Payoffs are such that  $u_i(a_1, a_2) = \min\{a_1, a_2\} - a_i \cdot d$ , where  $d$  is equal to 0.1 in one treatment and 0.9 in the other. Independent of  $d$ , any common effort level is a Nash equilibrium in this game. The high effort equilibrium is the efficient one. While the pure-strategy Nash equilibria are unaffected by this change in payoffs, GH’s experimental data show that agents exert lower effort when the cost is higher.

**Kreps Game.** The baseline and the modified games are described in the following table. The numbers in parenthesis represent the empirical distributions observed in the experiment:

Baseline:	<i>Left</i> (26)	<i>Middle</i> (8)	<i>Non Nash</i> (68)	<i>Right</i> (0)
<i>Top</i> (68)	200, 50	0, 45	10, 30	20, -250
<i>Bottom</i> (32)	0, -250	10, -100	30, 30	50, 40
Modified:	<i>Left</i> (24)	<i>Middle</i> (12)	<i>Non Nash</i> (64)	<i>Right</i> (0)
<i>Top</i> (84)	500, 350	300, 345	310, 330	320, 50
<i>Bottom</i> (16)	300, 50	310, 200	330, 330	350, 340

Notice that the modified game is obtained from the first simply by adding a constant of 300 to every payoff. This game has two pure-strategy equilibria,  $(Top, Left)$  and  $(Bottom, Right)$ , and one mixed-strategy equilibrium in which row randomizes between *Top* and *Bottom* and column randomizes between *Left* and *Middle*. Clearly, the change in payoff here only affects the weights in the mixed-equilibrium, not the pure equilibria. Yet, a majority of column play the Non-Nash action, and the change in payoffs only seems to affect the Row players.

The results of these experiments stand in sharp contrast with standard equilibrium concepts, and seem to preclude a game theoretic explanation. For instance, in the Kreps’ game most of individuals play an action that is inconsistent with Nash equilibrium; in the Traveler’s Dilemma and in the Effort Coordination game, changes in payoffs that do not affect the set of equilibria are associated to significant changes in behavior; in the Matching Pennies, Nash equilibrium seems to work only by coincidence in the baseline treatment, but asymmetries in payoffs determine major departures from the equilibrium predictions. Other heuristics rules, such as assuming that individuals play according to their ‘maximin’ strategy, or based on risk or loss aversion, may explain the behavior observed in some games, but not in others.<sup>10</sup> Yet, many of the anomalous

<sup>9</sup>Besides the data summarized in the matrices, GH also report (ibid., p.1408) that 64 percent of row players and 76 of the column players play *H* in the contradiction treatment ( $x = 400$ ). These data however are inconsistent: if 76 percent of column play *H* and 32 percent of observations are  $(H, H)$ , then cell  $(L, H)$  must receive a weight of 44. Since  $(L, L)$  is observed 16 percent of the times, it follows that at least 60 percent of row played *L*, whereas GH report that 64 percent of Row played *H*. We will return to this point in Section 5.1.2.

<sup>10</sup>As GH point out, the ‘maximin’ criterion may explain the behavior observed in the Kreps’ game, but is inconsistent with the evidence in the Matching Pennies and the Coordination games, and in any case it would entails unrealistic degrees of risk-aversion relative to what is commonly observed. Similarly, explanations based on loss-aversion would contradict the evidence from Kreps’ game, and losses are not a possibility in most of these games.

data are related to the nature of the incentives in an intuitive way, which suggests that it should be possible to develop formal models that account for these patterns of behavior. GH observe that simple models of level- $k$  reasoning may accommodate the experimental findings in the Traveler’s Dilemma, Kreps’ Game and in the Effort Coordination game

### 5.1.1 Five Little Treasures: A Unified Explanation

Throughout this section, we will consider the most stringent representation obtained in Section 3, in which the agent is particularly cautious about the validity of his current understanding of the game, implying that his his disposition to further think about the game is particularly strong. This corresponds to the representation provided in Theorem 3, where at each step the ‘value of reasoning’ for player  $i$ ,  $W_i(k)$ , is equal to the maximum difference between the payoff that the player could get if he chose the optimal action  $a_i^*$  and the payoff he would receive given his current action  $a_i^{k-1}$ , out of all the possible opponent’s actions. That is:

$$W_i(k) = \max_{(a_i, a_{-i}) \in A} u_i(a_i^*(a_{-i}), a_{-i}) - u_i(a_i^{k-1}, a_{-i}).^{11} \quad (13)$$

We will use this representation to contrast each of the ‘treasures’ with the corresponding ‘contradiction’. As discussed in Section 2.3, this presumes that each treasure and the corresponding contradictions belong to the same cognitive equivalence class. Given the nature of the payoff transformations involved in GH’s exercises, this assumption is consistent with criterion (C.1) for the cognitive partition (p. 8).

We begin by considering the games that GH observe could be explained by standard models of level- $k$  reasoning, though at the cost of ad hoc assumptions, which GH discard as not being particularly convincing (ibid., p. 1417):

“It is easy to verify that level one rationality also provides good predictions for both treasure and contradiction treatments in the traveler’s dilemma, the minimum-effort coordination game, and the Kreps game. There is evidence, however, that at least some subjects form more precise beliefs about others’ actions, possibly through higher levels of introspection.”

Consider the Minimum-Effort Coordination game first. Of all the treasures in GH, this seems to be the least suitable to being approached through some form of level- $k$  reasoning. On the one hand, any degenerate symmetric profile is a Nash equilibrium, hence if taken as an anchor it would generate the same profile at all iterations of reasoning process. On the other hand, there is no obvious way of approaching this game, hence no obvious specification of  $a^0$ . If, however, we follow GH and appeal to the principle of insufficient reason to set the anchor equal to the uniform distribution, then the level-1 levels of efforts with the low and high cost games are 164 and 116, respectively.

In the Traveler’s Dilemma, two specifications of  $a^0$  seem plausible and have been used in the literature. Namely, the anchor can be specified  $a_i^0 = 300$  or as the uniform distribution. In

---

<sup>11</sup>We focus on this representation for ease of exposition. The results that follow do not rely on this specific functional form. For instance, the results hold for the representation provided in Theorem 2 as well. That representation however involves a larger number of parameters, which would leave us more freedom in the calibration exercise. The parsimony of parameters in this representation, if anything, ties our hands more in the quantitative exercise.

the latter case, standard level- $k$  models roughly suffice to account for GH’s findings, through a mechanism similar to the one discussed for Minimum-Effort Coordination game. If  $R = 5$ , the best response to the uniform is  $a_i^1 = 290$ , and the path of reasoning is such that  $a_i^k = 291 - k$  for each  $k \in \mathbb{N}$ . If instead  $R = 180$ , then the best response to the uniform is 180, hence the shift in behavior may be explained by ‘level-1 rationality’ in both cases, merely due to a change in the best response.

While this is consistent with our model, we can also offer an alternative explanation. In particular, if  $a^0$  is specified such that  $a_i^0 = 300$  in the Traveler’s Dilemma, then  $a_i^k = 300 - k$  for each  $k \in \mathbb{N}$ . The change in behavior in the treasure and contradiction treatment in this case can be explained by the stronger ‘incentives to reason’ that the game provides when  $R$  is increased from 5 to 180. To see this, notice that (13) in this game is equal to  $\max\{299 - a_i^{k-1}, 2R - 2\}$ . Hence, with  $R = 5$ , we have for instance that  $W(k) = 8$  for all  $k = 1, \dots, 9$ . This implies that any individual for whom  $c(1) > 8$  would play 300. When  $R = 180$  instead the value of reasoning is such that  $W(k) = 358$  for all  $k < 120$ , while  $W(120) = 179$  (this is the number of steps in which the level- $k$  reasoning reaches the Nash equilibrium in this game.) Hence, anyone with a cost function such that  $c(1) > 8$  and  $c(k) < 180$  for all  $k$  would play 300 when  $R = 5$  and 180 when  $R = 180$ .<sup>12</sup> We also notice that any representation consistent with Theorem 1 would deliver qualitatively similar implications, because changing  $R$  from 5 to 180 increases the payoff differences of the game, hence the ‘incentives to reason’. This would suffice to conclude the individuals depth of reasoning would be larger in the high reward treatment.

The explanation based on ‘level-1 rationality’ has other important limitations. For the low reward treatment the level-1 action is 290, while most of the subjects actually play 300. Furthermore, even if this discrepancy is ignored, and the evidence explained in terms of ‘level-1’ rationality, it would remain the puzzle of why in this game such a high fraction of subjects plays according to level-1 rationality, whereas a classical finding of the literature on level- $k$  is that most individuals typically perform between two and three rounds of reasoning. The second explanation overcomes both limitations, endogenizing 300 precisely as the action that would be played, due to the low incentives to reason provided by the game payoffs when  $R = 5$ . We also find that our approach addresses the concern expressed in the quote above and acknowledges that individuals form more precise beliefs through higher levels of introspection.

We next turn to the Kreps’ Game, for which the results seem less intuitive than others in GH. This is perhaps because there is no obvious way of approaching this game, other than equilibrium play, but it is highly inconsistent with the observed behavior. The message of our theory for this treasure, however, is the simplest of all. Namely, since the payoff differences are unchanged in the baseline and modified game, any representation consistent with Theorem 1 (with Axiom 2, in fact) implies that whatever we observe in the baseline game should not change for the modified game. This is extremely close to what we observe, especially for the column players (less so for the row players, where the fraction playing *Top* goes from 68% to 83%). The distributions observed in the baseline game could of course be matched through appropriate choices of other parameters of the model, but the bite of our analysis is in the comparative statics, according to which there should be no difference in behavior between the two treatments.

---

<sup>12</sup>In Alaoui and Penta (2013), we argue that in games such as the Traveler’s Dilemma, where the agents may understand the inductive pattern in the chain of best responses, a sensible shape for the cost function is one in which only the first few steps are costly. Such a hump-shaped cost function, though not necessary for the results, would generate precisely this switch from 300 to 180 in the low and high reward treatment.

Besides overcoming the weakness of standard level- $k$  models to explain the games above, we show next that our model also explains the findings in the other treasures and contradictions, for which no clear explanation was provided: Matching Pennies and the Coordination Game with a Secure Option. We show that our model not only provides a qualitative explanation for GH’s findings, but also performs surprisingly well from a quantitative viewpoint. In particular, we perform a calibration exercise based on an extremely parsimonious specification of parameters. Holding such parameters constant across all the five treasures and contradictions, we show that the model’s prediction are surprisingly close to the empirical distributions observed in GH. The calibration exercise also illustrates how our model can be used for structural work on strategic thinking.

### 5.1.2 A Calibration Exercise

We assume that there are two types of players, one (strictly) more sophisticated than the other, respectively denoted by ‘high’ and ‘low’, and that the cost functions are strictly increasing in each of the treasures and finite for every  $k$ . We maintain that the distribution of types remains the same throughout all games and for both the row and column players, and let  $q_l$  denote the fraction of the low type. Lastly, while in some cases it would seem plausible that some level-zeros are more salient than others, we maintain throughout that anchors are uniformly distributed in both populations and for all games. This neutrality serves to illustrate that our assumptions are neither post-hoc nor arbitrary. The only parameter that we will use in our calibration will thus be the fraction of low types in the population,  $q_l$ .

**Matching Pennies.** Applying equation (13) to the matching pennies game, parameterized by  $x$ , the value of reasoning function takes the following shape:

$$W_1(k) = \begin{cases} 40 & \text{if } a_1^{k-1} = T \\ x - 40 & \text{if } a_1^{k-1} = B \end{cases} \quad (14)$$

$$W_2(k) = 40.$$

Since no pure-strategy action profile is a Nash equilibrium in this game, for any anchor  $a^0 \in A$  the reasoning process determines a cycle in the game. For instance, if the anchor of the row player is  $a^0 = (T, L)$ , then his sequence of actions  $(a_1^k)_{k=0,1,\dots} = (T, B, T, B, \dots)$ . Let  $c_l$  denote the (increasing) function of the low type, and consider the case  $x = 80$ . With  $x = 80$ , the  $W$  function is constant for both players. Given  $c_l$ , a low type may stop reasoning (hence play) at either action, depending on what his anchor is. If the anchors are distributed uniformly, it follows that the actions of the low types are distributed uniformly. For the same argument, 50 percent of the high types of population 1 will believe that the low type of the opponent plays  $L$  (resp.  $R$ ), and if  $q_l > 1/2$  this implies that they would play  $T$  ( $B$ ) irrespective of what their own cognitive bound is (hence, irrespective of their own cost function). Clearly, a similar argument applies to population 2. If instead  $q_l \leq 1/2$ , then the high types would play according to their own cognitive bound, which depends on the anchor. Once again, the uniformity assumption guarantees that the actions of the high types are uniformly distributed. Our model therefore predicts that, independent of  $q_l$ , actions are uniformly distributed in both populations. This prediction is remarkably close the

empirical results for  $x = 80$ . Notice, however, that the equilibrium distribution does not arise from equilibrium reasoning in this case, but from the assumption that anchors are distributed uniformly. In this sense, the model is consistent with the interpretation provided by GH, “In this context, *the Nash mixed-strategy prediction seems to work only by coincidence*, when the payoffs are symmetric.” (ibid., p. 1407)

Now, consider the case  $x = 320$ . The value of reasoning of the column players is not affected, hence the low types in population 2 will have the same behavior as in the case  $x = 80$ . To understand what happens to the low types of population 1, consider the following argument. Given any cost function  $c_l$ , individuals that (depending on their anchor) are unaffected, as  $W(k+1)$  does not change for them. Individuals who stopped reasoning at  $a_i^k = B$  when  $x = 80$  instead would now have a higher incentive to perform the next step of reasoning. If  $x$  is sufficiently high, then individuals such that  $a_i^k = B$  and  $a_i^{k+1} = T$  would perform one extra round of reasoning, but not two, whereas individuals for whom  $a_i^k = B$  and  $a_i^{k+1} = B$  would perform two extra rounds of reasoning, but not three (or one). In short, for any cost function  $c_l$ , This mechanism is true in general: it can be shown that, for any increasing cost function  $c_l$ , there exists  $\bar{x} > 80$  such that whenever  $x > \bar{x}$  all the low types playing  $T$  in the treasure won’t change, but all those playing  $B$  would switch to  $T$ . We assume that  $x = 320$  is high enough to affect the depth of reasoning. Hence, all low types in population 1, independent of  $a_0$ , would play  $T$  when  $x = 320$ .

We consider three cases next:

- If  $q_l > 1/2$ , then all high types of population 2 play  $R$ , best responding to the low types of population 1 playing  $T$ . It remains to consider what the high types of population 1 would do in this case. Given the depth of reasoning of the low types of population 2, the behavior that the high types of population 1 would expect from them depends on the anchor that they adopt. Given the uniformity assumption, 50 percent of high types in population 1 believe that the low types opponent play  $L$ , hence play  $T$ . The remaining 50 percent believes that the low types of population 1 play  $R$ , hence play  $B$ . It follows that, as long as  $q_l > 1/2$ , for  $x = 320$  we expect a fraction  $(q_l/2)$  of population 2 playing  $L$  and  $(\frac{1+q_l}{2})$  of population 1 playing  $T$ .
- If instead  $q_l \in (1/8, 1/2)$ , then we may have two cases: Since  $q_l > 1/8$ , the 50 percent of high types of population 1 that believe that the low types of population 2 play  $L$  will play  $T$ , and the 50 percent of high in population 2 that anticipate this will play  $R$ . The remaining 50 percent of high types in both populations will play according to their own cognitive bound. This will entail playing  $T$  for the high types of population 1, and (depending on the shape of the cost function) either  $R$  or  $L$  for the high types of population 2. It follows that, while the entire population 1 will play  $T$ , we may have two cases for population 2: either a uniform distribution over actions, or a distribution with weight  $(q_l/2)$  over  $L$ .
- If  $q_l < 1/8$ , then all high types of both populations will play according to their own cognitive bound. It follows that entire population 1 would play  $T$  and the actions of population 2 will be distributed uniformly.

If, based on the empirical results, we discard the explanations that suggest that population 2 would be uniformly distributed over  $L$  and  $R$ , we are left with two possibilities, both entailing that a fraction  $(q_l/2)$  of population 2 play  $L$ . If we choose  $q_l$  to match the empirical distributions, we

obtain  $q_l = .32$ , which falls precisely in the interval  $(1/8, 1/2)$ . The only explanation that appears consistent with the empirical distribution of population 2 therefore is the following, which requires that  $q_l \in (1/8, 1/2)$  and all agents in population 1 play  $T$ :

$x = 320$	L ( $q_l/2$ )	R ( $1 - q_l/2$ )
T (1)		
B (0)		

We next consider the case  $x = 44$ , maintaining that  $q_l \in (1/8, 1/2)$ . First, it is easy to show that for any increasing cost functions  $c_l, c_h$ , there exists  $x > 40$  sufficiently low that the both types of population 1 would choose  $B$  at their cognitive bound. Assuming that  $x = 44$  is such a ‘sufficiently low’ payoff, reasoning similar to the one above delivers the following results: all low types of population 1 play  $B$ , while the low types of population 2 are uniformly split; if  $q_l \in (1/8, 1/2)$ , the 50 percent of high types of population 1 that believe that the low types of population 2 play  $R$  will play  $B$ , and the 50 percent of high in population 2 that anticipate this will play  $L$ . The remaining 50 percent of high types in population 1 play according to their own cognitive bound, that is  $B$ . Since the high types in population 2 have the same cost function as the latter ones, but higher incentives, they would be able to anticipate this, and respond playing  $L$ . Hence, all high types in population 2 play  $L$ .

Summarizing, for  $q_l \in (1/8, 1/2)$  our findings for the three games are:

$x = 80$	L( $\frac{1}{2}$ )	R( $\frac{1}{2}$ )	$x = 320$	L( $\frac{q_l}{2}$ )	R( $1 - \frac{q_l}{2}$ )	$x = 44$	L( $1 - \frac{q_l}{2}$ )	R( $\frac{q_l}{2}$ )
T ( $\frac{1}{2}$ )			T (1)			T (0)		
B ( $\frac{1}{2}$ )			B (0)			B (1)		

The levels of  $q_l$  that maximize the fit in the two ‘contradictions’ are, respectively,  $q_l = .32$  for  $x = 320$  and  $q_l = .40$  for  $x = 44$ . Both these levels are consistent with the restriction  $q_l \in (1/8, 1/2)$ . The predicted distributions in the three treatments for these parameters are summarized next:

<i>Data</i>			$q_l = .32$			$q_l = .4$		
$x = 80$	L (48)	R (52)	$x = 80$	L (50)	R (50)	$x = 80$	L (50)	R (50)
T (48)			T (50)			T (50)		
B (52)			B (50)			B (50)		
$x = 320$	L (20)	R (80)	$x = 320$	L (16)	R (84)	$x = 320$	L (20)	R (80)
T (96)			T (100)			T (100)		
B (4)			B (0)			B (0)		
$x = 44$	L (80)	R (20)	$x = 44$	L (84)	R (16)	$x = 44$	L (80)	R (20)
T (2)			T (0)			T (0)		
B (98)			B (100)			B (100)		

Clearly, any particular choice is somewhat arbitrary, but all  $q_l \in [0.32, 0.40]$  perform remarkably well across all three treatments.

**Coordination Game with a Secure Outside Option.** This game has two pure-strategy Nash equilibria,  $(L, L)$  and  $(H, H)$ , which are not affected by the value of  $x$ . Hence, anchors equal to

$(L, L)$  or  $(H, H)$  would generate a path of reasoning in which respectively  $L$  or  $H$  is repeated. Anchors  $(L, H)$  or  $(H, L)$  determine a cycle alternating between  $H$  and  $L$ , also independent on the value of  $x$ . The paths generated by anchors that involve  $S$  instead vary with the value of  $x$ . Consider anchor  $a^0 = (L, S)$  first: Given  $a_2^0 = S$  player 1 is indifferent between  $L$  and  $H$  if  $x = 0$ , hence  $a_1^1$  is distributed uniformly over  $L$  and  $H$ ;  $L$  instead is the only best response to  $a_1^0 = L$ , hence  $a_2^1 = L$ . Given that  $a_1^1$  is the uniform distribution over  $H$  and  $L$ , player 2 best responds with  $a_2^2 = H$ , while the best response to  $a_2^1 = L$  is  $a_1^2 = L$ . Hence  $a^2 = (L, H)$ . Given this, further iterations determine a cycle that alternates between  $H$  and  $L$ . When  $x = 400$  instead the best response to  $S$  is  $L$ , hence  $a^1 = (L, L)$  and the rest of the path of reasoning yields  $L$  at every step. The case in which the anchor is  $a^0 = (H, S)$  is symmetric: the path yields a cycle for  $x = 400$ , and  $H$  for all  $k \geq 2$  for  $x = 0$ .

Applying equation (13) to this game, with payoffs parameterized by  $x$ , we obtain the following value of reasoning functions:

$$W_1(k) = \begin{cases} 180 & \text{if } a_1^{k-1} = L \\ \max\{90, x\} & \text{if } a_1^{k-1} = H \end{cases} \quad (15)$$

$$W_2(k) = \begin{cases} 90 & \text{if } a_2^{k-1} = H \\ 180 & \text{if } a_2^{k-1} = L \\ 140 & \text{if } a_2^{k-1} = S \end{cases} \quad (16)$$

Note that, being dominated, action  $S$  is never part of any path of reasoning (for any  $k > 0$ ). Nonetheless, it shapes player 1's incentives to reason, as an increase in  $x$  changes the value of doing a step of reasoning when player 1 is in a state in which action  $H$  is regarded as the most sophisticated. Also, similar to the asymmetric matching pennies games discussed above, any path in which agents cycle between action  $L$  and action  $H$  induce a  $W$  function that alternates between 90 and 180. Whether the spikes are associated to odd or even  $k$ 's depends on the anchor. When  $x = 400$ , the incentives to reason do not change for player 2, but  $W_1$  changes alternating between 180 and 400: the 'spikes' at 400 replace what would be 'troughs' at 90 with  $x = 0$ .

The experimental results showed that 96 percent of player 1 and 84 percent of player 2 played  $H$  when  $x = 0$ . One possible explanation is that in the baseline coordination game the efficient equilibrium is sufficiently focal that most individuals approach the game with  $a^0 = (H, H)$  as an anchor. The change in behavior observed when  $x = 400$  could then only be explained by arguing that this payoff transformation changes the way the agents approach the game. While we think this is a plausible explanation, we maintain the assumptions that anchors are uniformly distributed, and explore to what extent the mere change in incentives may explain the observed variation in behavior, independent of the possible change in the anchors.

Under the assumption that anchors are uniformly distributed, the only way that such a strong coordination on  $H$  can be explained is by assuming that the 'spikes' and 'troughs' determined alternating between 180 and 90 are already sufficiently pronounced that the low types involved in a reasoning process that determines a cycle stop their reasoning at  $H$ . Hence, with  $x = 0$ , agents that approach the game with anchors  $a^0 = (L, L)$  play  $L$ , all others play  $H$  (because they either settle on a constant  $H$ , as in  $a^0 = (H, S), (H, H)$ , or they determine a cycle, as in

$a^0 = (H, L), (L, H), (L, S)$ :

$x = 0$	L (1/6)	H (5/6)	S (0)
L (1/6)			
H (5/6)			

$x = 0$	L	H (84)	S
L			
H (96)			

We next consider the case  $x = 400$ , maintaining the assumption that the anchors are uniformly distributed. For the same reasons discussed above, for any pair of (increasing) cost functions  $c_l, c_h$ , there is  $x$  sufficiently high that all low types of population 1 with a reasoning process that involves a cycle stop at  $L$ . If  $q_l < 2/3$ , however, this is not enough to induce the high types of population 2 to play  $L$  as well. Hence, if we insist on the interval  $[0.32, 0.40]$  calibrated above, both the low and the high types in population 2 play according to their cognitive bound. Since the incentives to reason were not affected by the change in  $x$  for these individuals, the assumptions above entail that they play  $H$ . Hence, in population 2, all individuals with anchors  $a^0 \neq (L, L), (L, S)$  play  $H$ , the others play  $L$ . It remains to consider the high types of population 1. Since with  $x = 400$  they have stronger incentives to reason than the high types of population 2, any of these types involved in a cycle anticipates that both types of population 2 would play  $H$ , hence they respond with  $H$ . Thus, in population 1, only the individuals whose anchor is  $a^0 = (H, H)$  and the high types with anchors  $a^0 \neq (L, L), (L, S)$  play  $H$ , that is a total of  $1/6 + \frac{(1-q_l)}{2}$ , or  $2/3 - q_l/2$ . The others play  $L$ . To determine the percentages of coordination in  $(L, L)$  and  $(H, H)$ , we assume independence in the distributions of play between the row and the column players.

Summarizing:

$x = 400$	L (1/3)	H (2/3)	S (0)
L ( $1/3 + q_l/2$ )			
H ( $2/3 - q_l/2$ )			

with  $q_l = .32$  calibrated from the matching pennies game:

$x = 400$	L (33)	H (67)	S (0)
L (49)	16		
H (51)		34	

$x = 400$	L	H	S
L	(16)		
H		(32)	

with  $q_l = .40$  calibrated from the matching pennies game:

$x = 400$	L (33)	H (67)	S (0)
L (53)	17		
H (47)		30	

$x = 400$	L	H	S
L	(16)		
H		(32)	

# Appendix

## A Proofs From Section 2

**Theorem 1** Under Axioms 1-6, for any cognitive equivalence class  $C \in \mathcal{C}$ , there exist functions  $W : \mathcal{U}_i \times \mathbb{N} \rightarrow \mathbb{R}$  and  $c : \mathbb{N} \rightarrow \mathbb{R}_+ \cup \{\infty\}$  such that for any game  $u \in C$  and for any  $k \in \mathbb{N}$ ,  $u_i^1 \succsim_k u_i^0$  if and only if  $W(u_i, k) \geq c(k)$ . Furthermore, for any  $k \in \mathbb{N}$ :

1. The cost of reasoning is non-negative:  $c(k) \geq 0$ .
2. For each  $u_i \in C$ ,  $W(u_i, k) \geq 0$ , and  $W(u_i, k) = 0$  whenever  $D(u_i, a_i^k) = \mathbf{0}$ .
3. For each  $u, u' \in C$ ,  $D(u_i, a_i^k) \geq D(u'_i, a_i^k)$  implies  $W(u_i, k) \geq W(u'_i, k)$ .

### Proof of Theorem 1:

**Step 1:** Fix  $C \in \mathcal{C}$ . For any  $k$ , let  $W(\cdot, k) : \mathcal{U}_i \rightarrow \mathbb{R}$  represent preferences  $\succsim^*$  defined as  $u_i \succsim^* v_i$  if and only if  $u_i - u_i^k \succsim_k^1 v_i - v_i^k$ . Preferences  $\succsim^*$  inherit the weak order and archimedean properties of  $\succsim_k^1$ , hence such a representation  $W(\cdot, k)$  exists. For some  $v_i$ , we normalize  $W$  so that  $W(v_i^k, k) = 0$ . By construction, and by monotonicity of  $\succsim_k^1$ ,  $W(\cdot, k)$  is increasing in  $D(u_i, a_i^k)$  and  $W(u_i, k) = 0$  whenever  $D(u_i, a_i^k) = \mathbf{0}$ . (This is so because  $u_i - u_i^k = \mathbf{0} = v_i^k - v_i^k$  whenever  $D(u_i, a_i^k) = \mathbf{0}$ , hence  $W(u_i, k) = W(v_i^k, k) = 0$ .) Furthermore, for any  $u \in C$ ,  $W(u_i, k) \geq W(u_i^k, k) = W(v_i^k, k)$ : This follows from Axiom 5 which implies that  $u_i \succsim_s^1 u_i^k$  for all  $u \in C$ , and since  $u \in C$  implies  $(u - u_i^k) \in C$  (by condition 2.1) Axiom 5 implies that  $u_i - u_i^k \succsim_k^1 u_i^k - u_i^k$  (this is because  $u_i^k - u_i^k = (u_i - u_i^k)^k \in \Lambda(u_i - u_i^k)$ ) that is  $u_i \succsim^* u_i^k$ . Hence  $W(u_i, k) \geq W(u_i^k, k) = 0$  for all  $u \in C$ .

**Step 2:** If for all  $u_i \in \mathcal{U}_i$ ,  $u_i^0 \succsim_k u_i^1$ , then set  $c(k) = \infty$ , and the rest of the theorem holds trivially. If not, let  $\hat{u}_i \in \mathcal{U}_i$  be such that  $\hat{u}_i^1 \succ_k \hat{u}_i^0$ .

**Step 3:** If  $u_i^1 \succsim_k u_i^0$ , Axiom 6 implies  $u_i^0 \sim_k u_i^{k,0}$  and  $2(\text{part S.1}) u_i^{k,0} \succsim_k u_i^{k,1}$ , hence (by transitivity, Axiom 1),  $u_i^1 \succsim_k u_i^{k,1}$ .

**Step 4:** We show next that there exists  $c \in \mathbb{R}_+$  such that  $(u_i^k + c, 1) \succ_k u_i^{k,0}$ : Suppose not, then  $u_i^1 \succsim_k u_i^0 \sim_k u_i^{k,0} \succsim_k (u_i^k + c, 1)$  for all  $c \in \mathbb{R}_+$ , hence  $u_i^1 \succsim_k (u_i^k + c, 1)$  for all  $c$ . The latter contradicts Axiom 4.M. Hence, overall  $(u_i^k + c, 1) \succ_k u_i^{k,0} \succsim_k u_i^{k,1}$ . This implies (by Axiom 4.C.2) that there exists  $c(k) \in [0, c)$  such that  $u_i^{k,0} \sim_k (u_i^k + c(k), 1)$ . Hence we have  $u_i^1 \succsim_k (u_i^k + c(k), 1)$ , equivalently written  $u_i \succsim_k^1 u_i^k + c(k)$ . Furthermore, since  $D(u_i) = D(u_i - u_i^k)$ , Axiom 2 (S.2) implies:

$$u_i^1 \succsim_k u_i^0 \text{ if and only if } (u_i - u_i^k, 1) \succsim_k (u_i - u_i^k, 0). \quad (17)$$

**Step 5:** Replacing  $u_i^1 \succsim_k u_i^0$  with  $(u_i - u_i^k, 1) \succsim_k (u_i - u_i^k, 0)$  in Steps 3-4, it follows that  $u_i^1 \succsim_k u_i^0$  if and only if  $(u_i - u_i^k, 1) \succsim_k \left( (u_i - u_i^k)^k + c(k), 1 \right)$ , where  $(u_i - u_i^k)^k$  denotes the game with payoffs  $(u_i - u_i^k)^k(a_i, a_{-i}) := u_i(a_i^k, a_{-i}) - u_i^k(a_i, a_{-i}) \equiv 0$  for every  $(a_i, a_{-i}) \in A$ . Hence, we conclude that  $u_i^1 \succsim_k u_i^0$  if and only if  $u_i - u_i^k \succsim_k^1 c(k)$ .

**Step 6:** For each  $u_i$ , let  $t^{u_i} \in \mathbb{R}$  be such that  $(u_i + t^{u_i}, 1) \sim_k u_i^0$ . Notice that  $t^{u_i} \geq 0$  if and only if  $u_i^0 \succsim_k u_i^1$ . Furthermore, by axiom 2.2,  $t^{u_i} = t^{u_i - u_i^k}$ , because  $D(u_i) = D(u_i - u_i^k)$ . Consider the following specification of the  $W$  function: for each  $u_i$ , let  $W(u_i, k) = c(k) - t^{u_i}$ . We need to show that this specification actually represents the  $\succsim^*$  preference system, and that,

consistent with Step 1,  $W(u_i^k, k) = 0$  for all  $u_i$ . The latter property is immediate. Next, notice that  $W(u_i, k) \geq W(v_i, k)$  if and only if  $t^{u_i} \leq t^{v_i}$ , but since  $t^{u_i} = t^{u_i - u_i^k}$ ,  $t^{u_i} \leq t^{v_i}$  if and only if  $t^{u_i - u_i^k} \leq t^{v_i - v_i^k}$  if and only if  $u_i - u_i^k \succsim_k^1 v_i - v_i^k$ , that is  $u_i \succsim^* v_i$ . ■

**Theorem 2** Under Axioms 1-6 and 7, for any cognitive equivalence class  $C \in \mathcal{C}$ , there exists  $c : \mathbb{N} \rightarrow \mathbb{R}_+$  and for any  $k \in \mathbb{N}$  a simple probability distribution  $p^k \in \Delta(M_i(C))$  such that for any game  $u \in C$ ,  $u_i^1 \succsim_k u_i^0$  if and only if

$$\sum_{\mu \in \Delta(A_{-i})} p^k(\mu) \sum_{a_{-i}} \mu(a_{-i}) [u_i(a_i^*(\mu), a_{-i}) - u_i(a_i^k, a_{-i})] \geq c(k), \quad (18)$$

where  $a_i^*(\mu) \in BR_i^u(\mu)$  for any  $\mu \in M_i(C)$ .

**Proof:** Given Theorem 1, we only need to show that for any  $k$ , the function  $W(\cdot, k) : \mathcal{U}_i \rightarrow \mathbb{R}$  in that representation has the form of the left-hand side in eq. 18. To this end, notice that with the addition of the independence axiom the preference system  $\succsim_k^1$  satisfies the conditions for the mixture space theorem for every  $s$ . Hence, for any  $k \in \mathbb{N}$ , there exists a function  $\mathcal{V}'(\cdot, k) : \mathcal{U}_i \rightarrow \mathbb{R}$  that represents  $\succsim_s^1$  and satisfies  $\mathcal{V}'(\alpha u_i + (1 - \alpha)v_i, k) = \alpha \mathcal{V}'(u_i, k) + (1 - \alpha) \mathcal{V}'(v_i, k)$  for all  $\alpha \in [0, 1]$  and  $u_i, v_i \in \mathcal{U}_i$ . Moreover,  $\mathcal{V}'(\cdot, k)$  is unique up to positive affine transformations. Because  $\mathcal{V}'(\cdot, k)$  is linear in  $u_i \in \mathcal{U}_i = \mathbb{R}^{|A|}$ , there exist  $(\rho'(a))_{a \in A} \in \mathbb{R}^{|A|}$  s.t.  $\mathcal{V}'(u_i, k) = \sum_{a \in A} \rho'(a) \cdot u_i(a)$ . By monotonicity,  $\rho'(a) \geq 0$  for each  $a$ , and we define  $\rho^k$  normalizing such weights in the unit simplex, so that  $\rho^k(a) = \rho'(a) / \sum_{a' \in A} \rho'(a')$ . Since this is a positive affine transformation,  $\mathcal{V}(u_i, k) = \sum_{a \in A} \rho^k(a) \cdot u_i(a)$  also represents  $\succsim_k^1$  by the uniqueness part of the mixture space theorem. For any such  $\rho^k$ , define  $p^k \in \Delta(A_i)$  and  $\mu = (\mu^{a_i})_{a_i} \in \Delta(A_{-i})^{A_i}$  as follows: for any  $a_i \in A_i$ , let  $p^k(a_i) = \sum_{a_{-i}} \rho^k(a_i, a_{-i})$  and define  $\mu^{a_i} \in \Delta(A_{-i})$  such that, for any  $a_{-i}$ ,

$$\mu^{a_i}(a_{-i}) = \begin{cases} \frac{\rho^k(a_i, a_{-i})}{p^k(a_i)} & \text{if } p^k(a_i) > 0 \\ \frac{1}{|A_{-i}|} & \text{otherwise} \end{cases}.$$

It is immediate to see that, for any  $(a_i, a_{-i}) \in A$ ,  $\rho^k(a_i, a_{-i}) = p^k(a_i) \cdot \mu^{a_i}(a_{-i})$ . Hence, without loss of generality we can represent  $\mathcal{V}_s^1$  as follows:

$$\mathcal{V}(u_i, k) = \sum_{a_i} p^k(a_i) \sum_{a_{-i}} \mu^{a_i}(a_{-i}) \cdot u_i(a_i, a_{-i}). \quad (19)$$

We show next that  $u_i^* \in C$ ,  $a_i \in BR_i(\mu^{a_i}; u_i^*)$  whenever there exists  $a_{-i} \in A_{-i}$  such that  $\rho(a_i, a_{-i}) > 0$ .

Suppose not. Then  $\exists \hat{a}_i$  s.t.  $\rho(\hat{a}_i, a_{-i}) > 0$  for some  $a_{-i}$  s.t.  $\hat{a}_i \notin BR_i(\mu^{\hat{a}_i}; u_i^*)$ . Then, let  $a_i^* \in BR_i(\mu^{\hat{a}_i}; u_i^*)$  and define the relabeling  $\lambda : A_i \rightarrow A_i$  so that

$$\lambda(a_i) = \begin{cases} a_i^* & \text{if } a_i = \hat{a}_i \\ a_i & \text{otherwise} \end{cases}$$

Furthermore, let  $u_i^\lambda : A \rightarrow \mathbb{R}$  be such that, for any  $(a_i, a_{-i}) \in A_i$ ,  $u_i^\lambda(a_i, a_{-i}) = u_i^*(\lambda(a_i), a_{-i})$ . Then,

$$\begin{aligned}
\mathcal{V}(u_i^\lambda, k) &= \sum_{a \in A} \rho^k(a) \cdot u_i^\lambda(a) \\
&= \sum_{(a_i, a_{-i}) : a_i \neq \hat{a}_i} \rho^k(a_i, a_{-i}) \cdot u_i^*(a_i, a_{-i}) + \sum_{a_{-i} \in A_{-i}} \rho^k(\hat{a}_i, a_{-i}) \cdot u_i^*(a_i^*, a_{-i}) \\
&> \sum_{(a_i, a_{-i}) : a_i \neq \hat{a}_i} \rho^k(a_i, a_{-i}) \cdot u_i^*(a_i, a_{-i}) + \sum_{a_{-i} \in A_{-i}} \rho^k(\hat{a}_i, a_{-i}) \cdot u_i^*(\hat{a}_i, a_{-i}) \\
&= \mathcal{V}(u_i^*, k).
\end{aligned}$$

Since, by construction,  $u_i^\lambda \in \Lambda(u_i^*)$ , this conclusion contradicts Axiom 5.

Hence, whenever  $u_i^* \in C$ ,

$$\mathcal{V}(u_i^*, k) = \sum_{a_i} p^k(a_i) \sum_{a_{-i}} \mu^{a_i}(a_{-i}) \cdot u_i^*(a_i^*(\mu), a_{-i}). \quad (20)$$

Now, notice that functional  $W$  in Theorem 1 represented the preference ordering  $\succsim^*$  defined as  $u_i \succsim^* v_i$  if and only if  $u_i - u_i^k \succsim_k^1 v_i - v_i^k$ . Hence, since  $\mathcal{V}(\cdot, k)$  represents  $\succsim_k^1$ , we have that  $u_i \succsim^* v_i$  if and only if  $\mathcal{V}(u_i - u_i^k, k) \geq \mathcal{V}(v_i - v_i^k, k)$ . That is, thanks to the introduction of Axiom 7 we can set  $W(u_i, k) = \mathcal{V}(u_i - u_i^k, k)$ . Since

$$\mathcal{V}(u_i - u_i^k, k) = \sum_{a_i} p(a_i) \sum_{a_{-i}} \mu^{a_i}(a_{-i}) \cdot [u_i(a_i^*(\mu^{a_i}), a_{-i}) - u_i(a_i^k, a_{-i})],$$

the representation follows from Theorem 1, noticing that  $\mathcal{V}(u_i^k - u_i^k, k) = 0$ , and that (by Theorem 1)  $u_i^1 \succsim_k u_i^0$  if and only if  $u_i - u_i^k \succsim_k^1 c(k)$ , and noticing that  $\mathcal{V}(c(k), k) = c(k)$ . ■

**Proposition 1** Under Axioms 1-7 and 8, the representation in Theorem 2 is such that, for each  $C \in \mathcal{C}$ , for each  $u \in C$  and for each  $k$ ,  $a_i^k \in BR^u(\hat{\mu}^k)$ .

**Proof:** Suppose not, i.e. there exists  $a_i'$  s.t.  $\sum_{a_{-i}} \hat{\mu}^k(a_{-i}) u_i(a_i', a_{-i}) > \sum_{a_{-i}} \hat{\mu}^k(a_{-i}) u_i(a_i^k, a_{-i})$ . Then, substituting the definition of  $\hat{\mu}^k$ , we obtain

$$\sum_{\mu \in \Delta(A_{-i})} p^k(\mu) \sum_{a_{-i}} \mu(a_{-i}) [u_i(a_i', a_{-i}) - u_i(a_i^k, a_{-i})] > 0$$

hence  $v_i' \succsim_k^1 u_i^k$  where  $v_i' \in \Lambda(u_i)$  is such that  $v_i(a_i, a_{-i}) = u_i(a_i', a_{-i})$  for all  $(a_i, a_{-i}) \in A$ . ■

**Lemma 1** Under Axioms 1-6,  $u_i - u_i^k \succsim_s^1 v_i - v_i^k$  if and only if  $v_i^1 \succsim_s v_i^0$  implies  $u_i^1 \succsim_s u_i^0$ .

**Proof:**

( $\Rightarrow$ ) Suppose that  $u_i - u_i^k \succsim_s^1 v_i - v_i^k$  and that  $v_i^1 \succsim_s v_i^0$ . Then, eq. (17) implies that  $v_i - v_i^k \succsim_s^1 c(k)$ , hence  $u_i - u_i^k \succsim_s^1 c(k)$  by transitivity, and  $u_i^1 \succsim_s u_i^0$  follows from eq. (17).

( $\Leftarrow$ ) Suppose that  $v_i^1 \succsim_s v_i^0$  (if not, the statement holds vacuously) and (as contrapositive) that  $u_i^0 \succ_s u_i^1$ . Then eq. (17) implies  $v_i - v_i^k \succsim_s^1 c(k) \succ_s^1 u_i - u_i^k$ , which contradicts  $u_i - u_i^k \succsim_s^1 v_i - v_i^k$ . ■

**Theorem 3:** Under Axioms 1-6 and 9, for any cognitive equivalence class  $C \in \mathcal{C}$ , there exist a function  $c : \mathbb{N} \rightarrow \mathbb{R}_+ \cup \{\infty\}$  such that for any game  $u \in C$  and for any  $k \in \mathbb{N}$ ,  $u_i^1 \succsim_k u_i^0$  if and

only if

$$\max_{a_{-i} \in A_{-i}} u_i(a_i^*(a_{-i}), a_{-i}) - u_i(a_i^k, a_{-i}) \geq c(k).$$

**Proof:** Given Theorem 1 and Lemma 1, we only need to show that  $D^* : \mathcal{U}_i \rightarrow \mathbb{R}$  defined as  $D^*(u_i) = \max_{a_{-i} \in A_{-i}} u_i(a_i^*(a_{-i}), a_{-i}) - u_i(a_i^k, a_{-i})$  represents the preference system  $\succsim_s^*$ . Suppose not, then there exist  $u_i, v_i \in \mathcal{U}_i$  such that  $D^*(u_i) > D^*(v_i)$  but  $v_i \succsim_s^* u_i$ , that is  $v_i - v_i^k \succsim_s^1 u_i - u_i^k$ . Now,

$$\hat{a}_{-i} \in \arg \max_{a_{-i} \in A_{-i}} u_i(a_i^*(a_{-i}), a_{-i}) - u_i(a_i^k, a_{-i})$$

and let  $\hat{u}_i$  be such that

$$\hat{u}_i(a_i, a_{-i}) = \begin{cases} u_i(a_i^k, \hat{a}_{-i}) & \text{if } a_i = a_i^k \\ u_i(a_i^*(\hat{a}_{-i}), \hat{a}_{-i}) & \text{otherwise} \end{cases}.$$

Notice that  $\hat{u}_i \in \Gamma(u_i)$ , and that  $(\hat{u}_i - \hat{u}_i^k)(a) = D^*(u_i)$  for all  $a \in A$ , hence  $(\hat{u}_i - \hat{u}_i^k) \gg (v_i - v_i^k)$ . Monotonicity of  $\succsim_s^1$  thus implies that  $\hat{u}_i - \hat{u}_i^k \succ_s^1 v_i - v_i^k$ , that is  $\hat{u}_i \succ_s^* v_i$ , whereas Axiom 9 implies that  $u_i \succ_s^* \hat{u}_i$ . Overall,  $\hat{u}_i \succ_s^* v_i \succ_s^* u_i \succ_s^* \hat{u}_i$ , a contradiction. ■

## References

1. Agranov, Marina, Elizabeth Potamites, Andrew Schotter and Chloe Tergiman. 2012. "Beliefs and Endogenous Cognitive Levels: An Experimental Study" *Games and Economic Behavior*, 75(2): 449-463.
2. Alaoui, Larbi and Antonio Penta. 2013. "Depth of Reasoning and Incentives", *mimeo*.
3. Basu, Kaushik, "The Traveler's Dilemma: Paradoxes of Rationality in Game Theory." *American Economic Review Papers and Proceedings*, 84(2): 391-395.
4. Bhatt, Meghana A., and Colin F. Camerer. 2005. "Self-referential Thinking and Equilibrium as States of Mind in Games: fMRI Evidence." *Games and Economic Behavior*, 52(2): 424-459.
5. Bhatt, Meghana A., Terry Lohrenz, Colin F. Camerer, and P. Read Montague. 2010. "Neural Signatures of Strategic Types in a Two-Person Bargaining Game." *Proceedings of the National Academy of Sciences*, 107(46): 19720-19725.
6. Bosch-Domènech, Antoni, Jose García-Montalvo, Rosemarie Nagel, and Albert Satorra. 2002. "One, Two, (Three), Infinity...: Newspaper and Lab Beauty-Contest Experiments." *American Economic Review*, 92(5), 1687-1701.
7. Camerer, Colin F., Teck-Hua Ho, and Juin Kuan Chong. 2004. "A Cognitive Hierarchy Model of Games." *Quarterly Journal of Economics*, 119(3): 861-898.
8. Capra, C. Monica, Jacob K. Goeree, Rosario Gomez, Charles A. Holt. 1999. "Anomalous Behavior in a Traveler's Dilemma?" *American Economic Review*, 89(3): 678-690.
9. Choi, Syungjoo. 2012. "A Cognitive Hierarchy Model of Learning in Networks." *Review of Economic Design*, 16: 215-250.
10. Coricelli, Giorgio, and Rosemarie Nagel. 2009. "Neural Correlates of Depth of Strategic Reasoning in Medial Prefrontal Cortex." *Proceedings of the National Academy of Sciences*, 106(23): 9163-9168.
11. Costa-Gomes, Miguel A., and Vincent P. Crawford. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *American Economic Review*, 96(5): 1737-1768.
12. Costa-Gomes, Miguel A., Vincent P. Crawford, and Bruno Broseta. 2001. "Cognition and Behavior in Normal-Form Games: An Experimental Study." *Econometrica*, 69(5): 1193-1235.
13. Crawford, Vincent P. 2003. "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions." *American Economic Review*, 93(1): 133-149.
14. Crawford, Vincent P., Miguel A. Costa-Gomes, and Nagore Iriberrí. 2013. "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications." *Journal of Economic Literature*, 51.

15. Crawford, Vincent P., and Nagore Iriberrí. 2007. "Level- $k$  Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?" *Econometrica*, 75(6): 1721-1770.
16. Gabaix, Xavier. "Game Theory with Sparsity-Based Bounded Rationality." *mimeo*
17. Goeree, Jacob K., and Charles A. Holt. 2001. "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions." *American Economic Review*, 91(5): 1402-1422.
18. Goeree, Jacob K., and Charles A. Holt. 2004. "A Model of Noisy Introspection." *Games and Economic Behavior*, 46(2): 365-382.
19. Grosskopf, Brit, and Rosemarie Nagel. 2008. "The Two-Person beauty contest." *Games and Economic Behavior*, 62(1): 93-99.
20. Kets, Willemien. 2012. "Bounded Reasoning and Higher-Order Uncertainty." *mimeo*, Kellogg School of Management, Northwestern University.
21. Kneeland, Terri. 2012. "Coordination under Limited Depth of Reasoning." *mimeo*, University of British Columbia.
22. Ho, Teck-Hua, Colin Camerer, and Keith Weigelt. 1998. "Iterated Dominance and Iterated Best Response in Experimental 'p-Beauty Contests'." *American Economic Review*, 88(4): 947-969.
23. Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *American Economic Review*, 85(5): 1313-1326.
24. Palacios-Huerta, Ignacio and Oscar Volij. 2009. "Field Centipedes." *American Economic Review*, 99(4): 1619-1635.
25. Stahl, Dale O., and Paul R. Wilson. 1994. "Experimental Evidence on Players' Models of Other Players." *Journal of Economic Behavior and Organization*, 25(3): 309-327.
26. Stahl, Dale O., and Paul R. Wilson. 1995. "On Players' Models of Other Players: Theory and Experimental Evidence." *Games and Economic Behavior*, 10(1): 218-254.
27. Strzalecki, Tomasz. 2010. "Depth of Reasoning and Higher-Order Beliefs." Harvard Institute of Economic Research Discussion Paper Number 2184.