University of Leicester

Department of Economics

Evidential equilibria: Heuristics and biases in static games

Sanjit Dhami, University of Leicester

Ali al-Nowaihi, University of Leicester

Working Paper No. 13/25

November 2013

Evidential equilibria: Heuristics and biases in static games

Sanjit Dhami* **Ali al-Nowaihi†**

07 November 2013

Abstract

Standard equilibrium concepts in game theory find it difficult to explain the empirical evidence in a large number of static games such as prisoners' dilemma, voting, public goods, oligopoly, etc. Under uncertainty about what others will do in one-shot games of complete and incomplete information, evidence suggests that people often use *evidential reasoning* (ER), i.e., they assign *diagnostic significance* to their own actions in forming beliefs about the actions of other like-minded players. This is best viewed as a heuristic or bias relative to the standard approach. We provide a formal theoretical framework that incorporates ER into static games by proposing *evidential games* and the relevant solution concept- *evidential equilibrium* (EE). We derive the relation between a Nash equilibrium and an EE. We also apply EE to several common games including the prisoners' dilemma and oligopoly games.

Keywords: Evidential reasoning; causal reasoning; evidential games; social projection functions; ingroups and outgroups; evidential equilibria and consistent evidential equilibria; Nash equilibria; the prisoners' dilemma and oligopoly games; common knowledge and epistemic foundations.

*Corresponding author. Department of Economics, University Road, University of Leicester, Leicester LE1 7RH, UK. Phone: +44-116-2522086. Fax: +44-116-2522908. E-mail: Sanjit.Dhami@le.ac.uk.

†Department of Economics, University of Leicester, University Road, Leicester. LE1 7RH, UK. Phone: +44-116-2522898. Fax: +44-116-2522908. E-mail: aa10@le.ac.uk.

1 Introduction

A considerable body of evidence shows that the predictions of the standard equilibrium concepts in game theory are not borne out by a significant fraction of experimental subjects.[1] In this paper, we are interested in static games of complete and incomplete information such as a one-shot prisoners' dilemma game, a one-shot voting game, a one-shot public goods game, and duopoly games such as Cournot, Bertrand or Stackelberg games etc.

Let us briefly note the nature of the violations in some static games that we are interested in. A more detailed treatment and references are given in Section 2. More than half the subjects in the prisoners' dilemma game play the dominated action 'cooperate'. Voters vote in elections when it is clear to them that they will not be pivotal. Under traditional preferences, if there is a cost to voting, voting is a dominated action. The dominant action in public good games is to free-ride, yet we observe high levels of contributions initially that decay over time. Furthermore, if we combine *like-minded* players in public good games, based on the level of contributions, then we can elicit near first-best levels of contributions. In experiments with two players who must simultaneously choose outputs, the Cournot prediction is played at best by approximately only a third of the players, while monopoly, Stackelberg and even the competitive outcomes are quite common. The evidence from these and similar games suggests the following stylized facts that any reasonable theory of static games may aspire to explain.

S1. A significant fraction of players behave in a manner that is consistent with the predictions of classical game theory. For instance, many players defect in a prisoners' dilemma game, many people abstain from voting, many people contribute very low amounts or none at all in the early rounds of public goods games.

S2. An even larger fraction of players violate the predictions of classical game theory and they often seem to behave non-strategically. In particular, in certain games, they play dominated actions. But these dominated actions jointly lead to higher payoffs, e.g., both players cooperating in prisoners' dilemma games. These findings are not restricted to the games that we consider in this paper but are fairly robust in static games. For instance, in a second price common value auction, Ivanov et al. (2010) find that more than 30% of their subjects choose dominated bids. Controlling for lack of attention, misconception or insufficient incentives, de Sousa et al. (2012) find in a p-beauty contest with chess players that half of their experimental subjects behave in a non-strategic manner.[2]

[1] See, Camerer (2003) for a book length treatment.

[2] The authors argue that (p.3): "Our results overall suggest that the existence of non-strategic players in one-shot games is a robust feature. Rather than disregarding non-strategic players as noise, we believe

S3. In environments where players think that they are playing with other *like-minded* players they *impute diagnostic significance to their own actions when forming beliefs about the actions of others.*[3] For instance, those who cooperate (respectively defect) in prisoners' dilemma games think that the vast majority of other players will also cooperate (respectively defect). Similarly, despite publicly available information on election polls, those who vote Democrat (respectively Republican) in the US Presidential elections believe that a significant majority of other voters will also vote Democrat (respectively Republican).

Each of the static games that we consider have an exceedingly simple structure. Hence, we believe that the anomalies relative to the classical predictions are less likely to arise from mistakenly playing the incorrect action.[4] Rather, we believe that the formation of individual beliefs about how their opponents will play is the critical element.

1.1 Epistemic conditions

The Nash equilibrium for static games of complete information does not impose any restrictions on the beliefs of the players. On the other hand, the epistemic conditions that are known to imply a Nash equilibrium turn out to be extremely demanding. Aumann and Brandenburger (1995) proved the following result. Suppose that we have mutual knowledge of the payoffs, mutual knowledge of rationality, common knowledge of the beliefs (or conjectures) each player has about the others, and common priors. Then the common conjectures about each player j by the other players agree and constitute a Nash equilibrium of the game. Polak (1998) showed that if mutual knowledge of the payoffs is strengthened to common knowledge of payoffs then the Aumann-Brandenburger conditions imply common knowledge of rationality. One may ask how players acquire such information. A long history of experiments show that players do not acquire such information through repeated play of the game; see Colman et al. (2010). Furthermore, insofar as the Nash equilibrium is often violated in experiments (stylized fact S2), these epistemic conditions cannot empirically hold either.[5]

One of the leading belief-based behavioral models that has been quite successful in the explanation of the results of many static games has been the *Level-k model.* For instance,

that they should be considered as one of the main empirical regularities found in situations in which economic agents are confronted with new situations." Since the experimental subjects are chess players, experience from one domain does not necessarily transfer into another domain.

[3]We state some of the key features of stylized fact S3 here. A more detailed discussion with reference to the sources can be found in Section 2.

[4]For this reason our focus is not on behavioral alternatives such as *quantal response equilibrium* (QRE) in which players play a *noisy best response* but otherwise have consistent beliefs.

[5]Recent research offers an unsettled view of the epistemic foundations; see for instance, Gintis (2012).

it makes very specific predictions for the p-beauty contest game that are supported by the evidence; see, for instance, Camerer (2003). In Level-k models players play a best response, yet their beliefs may not be justified in equilibrium, hence, these models belong to the more general class of *disequilibrium in belief* models. Level-k models have been very successful in other contexts too, e.g., explaining the salience of focal points and in explaining the results from auctions.[6] However, under Level-k models, one should not observe the play of dominated strategies by any level-k player, $k > 1$, which makes it difficult to account for stylized fact S2.

1.2 Evidential reasoning (ER)

In the spirit of belief-based models, we examine the implications for static games of a particular kind of reasoning, *evidential reasoning* (ER), which is also sometimes known as *social projection*. ER is underpinned by a great deal of psychological evidence. When players who use ER are uncertain of the actions of other *like-minded players*, they assign *diagnostic significance* to their own actions in forming inferences about how others may behave. This might arise because there is true uncertainty in inferring what others will do in static games, or because we are simply hard-wired to use ER, or because we need to economize on scare cognitive resources, etc. ER may not apply to players who are not viewed as like-minded. Following the psychological literature, we may refer to the reasoning used in classical game theory as *causal reasoning*. A decision maker who uses causal reasoning assigns no diagnostic value to his own actions, unlike evidential reasoning.

Robbins and Krueger (2005) describe evidential reasoning thus: "Using their own disposition or preferences as data, people can make quick predictions of what others are like or what they are likely to do". In a recent survey, Krueger (2007) writes "The concept of social projection is once again generating vigorous theory development and empirical research ... social projection is among the simplest, oldest, and arguably most central concepts of the field".

We now consider some initial examples of ER.

Example 1 *(Grafstein, 1991): Suppose that an individual, A, would like to meet a colleague, B, who lives in the same area, on the way to work. B is known to be similar in personality and tastes but it is not known what route B will take to work. Then A may use her own choice of route (e.g., a 'scenic route' or a 'quite route' or the 'shortest route' etc.) as the best guess of what route B may take. Thus, one's own actions may have diagnostic significance in inferring the actions of others. By choosing a particular route, A cannot influence the route that B takes, nor does A believe so.*

[6] See, for instance, Crawford and Iriberri (2007).

Example 2 *(Prisoners' dilemma game): Consider the one-shot prisoners' dilemma game. A player who uses causal reasoning will reason as follows "My opponent is rational and uses causal reasoning. To defect is a strictly dominant strategy for me as well as my opponent, so I will defect". On the other hand, if a player who uses evidential reasoning perceives his opponent as like-minded, he may reason as follows "But if each of us defects, then we will be in a worse situation than had we cooperated. The other player is like-minded, so he will also realize that mutual cooperation is better for both of us than mutual defection. So I will cooperate". Whether players use such reasoning is a matter of empirical evidence. Indeed the evidence from prisoners' dilemma game (see Section 2.2) is consistent with a mixture of players who use causal and evidential reasoning. Furthermore, players who choose to play the action cooperate (defect) estimate with high probability that others will cooperate (defect), which is directly consistent with ER. The cooperative outcome in the prisoner's dilemma game under ER shows that welfare under ER may be higher than under causal reasoning.*

Example 3 *(False consensus effect): Ross et al. (1977) asked subjects if they would walk around a university campus wearing a sandwich board that said "REPENT". Those who agreed to do so, estimated that 63.5% of their peers would also agree to do so, while those who refused expected 76.7% of their peers to also refuse. Clearly these fractions add up to more than one and so cannot constitute consistent beliefs. This evidence is in line with subjects using evidential reasoning to impute diagnostic value to their actions in forming beliefs about the likely actions of other like-minded people (the student population in the university in this case). This is an example of the false consensus effect.*

Interestingly, people who use evidential reasoning are not aware of using it despite their behavior being obviously consistent with evidential reasoning. Evidential reasoning appears to arise as an *automatic* response, rather than a *deliberate* response, i.e., it does not require awareness, effort or intention. Evidence supporting this view comes from experiments which show that evidential reasoning was not hampered by cognitive load or time required to complete an action; see Krueger (2007). Furthermore, other evidence, also reported in Krueger (2007), suggests that considerable cognitive effort is required to suspend evidential reasoning. The evidence from Acevedo and Krueger (2005) indicates that evidential reasoning applies to human-human interaction but not to human-nonhuman interaction. Another feature of evidential reasoning is that individuals continue to behave in a self interested manner.

Economists accustomed to traditional notions of rationality will find evidential reasoning to be less than fully rational. And they are right. It is best, therefore, to view ER as a *heuristic or bias*. However, the relevant issue is whether the evidence on human behavior supports ER. Our own reading of the evidence is that the results of experimental games are best explained by a mixture of people who use evidential and causal reasoning. Our framework allows for such a mixture. At the same time it appears that there is no single theory that is capable of explaining the evidence from all static games.

In static games of complete information, when, for whatever reason, there is uncertainty about what others will do, evidential reasoning enables the creation of a set of beliefs about the likely actions of other players. As shown in the three examples above, the evidence is consistent with players who follow ER and, thus, believe that other like-minded players will follow the same actions that they themselves find desirable to follow. They then maximize their utility, conditional on these beliefs even though these beliefs may not turn out to be justified ex-post, as, say, in Level-k models[7]. This converts an essentially *strategic situation* to a *decision-theoretic problem*.

The motivation for our approach may be given with an analogy. Starting with the earliest experiments by Maurice Allais, the evidence from a very large number of experiments turned out to be inconsistent with the predictions of models of formal decision theory.[8] In the early 1970's Daniel Kahneman and Amos Tversky proposed an alternative approach, the *heuristics and biases* approach. In this approach, instead of optimizing in the classical sense, decision makers follow simple heuristics that are *fast* and *frugal* in the use of information. The heuristics often work well but sometimes lead to errors relative to a strictly optimizing behaviour.[9] Thus our proposed evidential equilibrium concept contributes to providing a heuristics and biases approach to strategic interaction. Indeed, many other alternatives to a Nash equilibrium, such as Level-k models, may also be viewed in a similar manner.

1.3 Structure of the paper

Further examples and a more detailed consideration of ER is given in Section 2 from many contexts, such as prisoners' dilemma, public goods, voting, etc. Section 3, *Evidential Games*, gives a formal treatment of evidential reasoning and proposes several concepts that we will find useful in the rest of the paper. An *evidential game* is simply a game where players use evidential reasoning. An *evidential equilibrium* is one where each player chooses to optimize given his beliefs about the behavior of the other players (inferred from his own behavior in accordance with evidential reasoning). A *consistent evidential equilibrium* is an evidential equilibrium where beliefs turn out to be correct although we,

[7] We consider both cases under the rubric of ER where beliefs are justified ex-post and where they are not.

[8] See, for instance, Kahneman and Tversky (2000) for the details of many of these results.

[9] A substantial literature developed subsequently that identified a rich range of heuristics and generated evidence that they are used by human subjects. See, for instance Kahneman et al. (1982) and the Nobel Lecture in Kahneman (2003).

equally, allow for *inconsistent evidential equilibria*. Our formulation of evidential reasoning yields causal reasoning, the dominant reasoning assumed in the traditional framework in economics, as a special case. If players use causal reasoning then a consistent evidential equilibrium corresponds to a Nash equilibrium in the ordinary sense; this is dealt with in more detail in subsection 3.5.

Sections 4 and 5 apply the theory developed in Section 3 to important games in economics and the social sciences. These are the prisoners' dilemma game (Section 4) and oligopoly games (Section 5).We conclude that the evidence from these games is more supportive of a mixture of evidential reasoning and causal reasoning.

So far, the discussion has been about static games of *complete* information. Section 6 extends the analysis of evidential reasoning to static games of *incomplete* information. For such games, we give a natural extension of an evidential equilibrium.

Section 7 concludes.

2 Evidential reasoning

We now offer further discussion and some concrete contexts in which evidential reasoning has been studied.

2.1 Calvinism and the development of capitalism

Consider the following example from Quattrone and Tversky (1984). According to the Calvinist doctrine of *predestination*, those who are to be saved have been chosen by God at the beginning of time, and nothing that one can do will lead to salvation unless one has been chosen. Although one cannot increase the chance of salvation by good works, one can produce *diagnostic evidence* of having been chosen by engaging in acts of piety, devotion to duty, hard work and self denial. According to Max Weber, this is exactly how millions of people responded to the Calvinist doctrine and why capitalism developed more quickly in Protestant rather than Catholic countries; see for example, Nozick (1993).[10]

2.2 Why is there so much cooperation in the prisoners' dilemma game?

The static prisoners' dilemma is a well known game in the social sciences in which each of the two players can either cooperate (C) or defect (D). The payoffs from each action to the row and the column player are shown in Table 1. In each cell, the first payoff is to the row player and the second to the column player and $x > 2$.

[10] We are grateful to Andrew Colman for drawing our attention to this example.

	C	D
C	$2,2$	$0,x$
D	$x,0$	$1,1$

Table 1: The Prisoner's Dilemma Game

Defection is a strictly dominant strategy. Hence, a player using causal reasoning should defect. However, experimental evidence indicates high cooperation rates. Rapoport (1988) finds cooperation rates of 50% in the prisoners dilemma game. In a study of the prisoners' dilemma based on high stakes outcomes from a British TV show called Goldenballs, Darai and Grätz (2010) finds cooperation rates of 55% for stakes above £500 and cooperation rates of 74% for stakes below this level. Zhong et al. (2007) show that the cooperation rates in prisoners' dilemma studies go up to 60% when positive labels are used (such as a "cooperative game" rather than a "prisoners' dilemma game"). When purely generic labels are used (such as C and D) then the cooperation rates are about 50%. In similar games, such as the one-shot public good contributions game, one also observes high cooperation rate; see for instance, Dawes and Thaler (1988).

In a novel study, Khadjavi and Lange (2013) find that actual prison inmates cooperate far more than the student population. In particular, while the cooperation rates among students playing the static prisoners' dilemma game is 37%, the cooperation rate among prison inmates is 56%.

2.2.1 Can cooperation in a prisoners' dilemma game be explained by existing theories?

Can the propensity to cooperate also be explained by *other regarding preferences*? In the standard model, individuals derive utility purely from their own levels of consumption (or *selfish preferences*). Evidence indicates, however, that, in addition to purely selfish considerations, individuals may also exhibit *altruism* and *envy*. Consider, for instance, *other regarding preferences* as in the model of Fehr and Schmidt (1999) (FS preferences, for short) that has been axiomatically founded and can explain the results of a very wide range of experimental games. Suppose that we have N individuals and n different income classes, where the typical income $y_j \in \mathbf{Y} = \{y_1 < y_2 < ... < y_n\}$. The proportion of individuals in the income class $y_j \in \mathbf{Y}$ is $p_j \geq 0$ and $\sum_{j=1}^{n} p_j = 1$. Then the FS utility function of an individual with income $y_j \in \mathbf{Y}$ is given by

$$U(y_j) = y_j - \beta \sum_{i=1}^{j-1} p_i (y_j - y_i) - \alpha \sum_{k=j+1}^{n} p_k (y_k - y_j), \alpha \geq 0, 0 \leq \beta < 1. \tag{1}$$

Individuals with FS preferences care for their own payoffs as under selfish preferences. But they also derive disutility from being ahead of others (altruism) and from being behind others (envy). $\beta \geq 0$ and $\alpha \geq 0$ are sometimes known as the parameters of, respectively, advantageous and disadvantageous inequity. When $\alpha = \beta = 0$ we have purely selfish preferences. Evidence indicates that disadvantageous inequity is more important than advantageous inequity ($\alpha > \beta$) and one never benefits by throwing away one's own income ($\beta < 1$).

Let us apply FS preferences to the prisoners' dilemma game. Suppose that half the players are row players and half are column players. The FS utility to any player from the strategy profiles (C, C) and (D, D) respectively is 2 and 1. When both players have FS preferences, is (C, C) a Nash equilibrium? Given that the column player plays C, the FS utility of the row player from defecting is $x - \frac{\beta}{2}x$. Thus, the row player will optimally choose to cooperate if $2 > x\left(1 - \frac{\beta}{2}\right)$ or equivalently, $\beta > 2 - \frac{4}{x}$ where $x > 2$. As x increases, the lower bound on β that is required to sustain cooperation increases. For instance, for $x = 3$ the lower bound is 0.67 and for $x = 5$ the lower bound is 1.2. By contrast, the empirical estimates of β are quite low. Fehr and Schmidt (1999) find that 60% of their subjects have $\beta \leq 0.25$ and 90% have $\beta \leq 0.60$. Fehr and Schmidt (2010) report some updated estimates from three other papers related to contractual interaction among agents. In these estimates, for 60% of the population, $\beta = 0$ and for 40% of the population, $\beta = 0.60$. These estimates suggest that it is hard to sustain cooperation in prisoners' dilemma games based on other regarding preferences.

The evidence also shows that *pure altruism/warm glow* cannot explain the evidence for cooperation in prisoners' dilemma games either. Cooper et al. (1996) assume that altruism takes the form of the row player deriving warm glow from the act of cooperation. Hence, Table 1 is modified to Table 2 where $\delta > 0$ captures warm glow to both players from cooperating.

	C	D
C	$2+\delta, 2+\delta$	$0+\delta, x$
D	$x, 0+\delta$	$1, 1$

Table 2: The Prisoner's Dilemma Game in the Presence of Warm Glow

If $\delta > x - 2$ then (C, C) becomes a Nash equilibrium and if $\delta > \min\{x - 2, 1\}$ then (C, C) is the unique pure strategy Nash equilibrium. If for a player $\delta > \min\{x - 2, 1\}$ then cooperation becomes a dominant strategy for each player; such players are termed as *altruists*. Players are randomly matched with each other over 20 rounds to eliminate repeated game effects. Cooper et al. (1996) find that 38% of the players in the first 10 rounds are altruists and 22% players in the last 10 rounds are altruists. However, altruism is not able to explain the extent of cooperation in the authors' experiments. Further, altruism is not able to explain the decay in cooperation. These results also contradict the theoretical result that cooperation can be sustained over a finite horizon if players perceive a small degree of irrationality about their opponents (see, for instance, Kreps et al. (1982)). Thus, neither of the two leading explanations in game theory account for the observed levels of cooperation in the one-shot prisoners' dilemma game.

Since the action C is a dominated strategy it is not *rationalizable* and cannot be supported in a *correlated equilibrium* either. In Level-k models, any player with level-k, $k \geq 1$ will also never play a dominated strategy.

For these reasons, the explanation of cooperative behavior in prisoners' dilemma games is an open problem in economics and social science.

2.2.2 Explaining cooperation in prisoners' dilemma games by evidential reasoning

Lewis (1979) used evidential reasoning to explain the unexpected levels of cooperation in the one-shot prisoners' dilemma game. Mutual cooperation is better than mutual defection. If players use evidential reasoning, they may take their own preference for mutual cooperation as diagnostic evidence that their rival also has a preference for mutual cooperation, in which case both players are more likely to cooperate. These views are borne out by the evidence. Cooperators (those who play C) believe that the probability of other players cooperating is between 0.6 and 0.7. Similarly, players who defect (those who play D) believe that other players will defect with probabilities between 0.6 to 0.7; see Krueger (2007). Darai and Grätz (2010) find that pre-play promises of players to cooperate with each other and shake hands on it, increases the cooperation rates. In classical theory this should have no effect, but evidential reasoning may provide one possible explanation. Such actions may increase the likelihood, in the minds of players, that they are facing like-minded players, hence, they cooperate more.

Rapoport (1966, p. 139-41) argued that each player takes his own belief that rational players should prefer the cooperative outcome as evidence that similarly rational players will also cooperate[11]. This is similar to evidential reasoning. Letting $x = 3$ in Table 1, this is how the row player is envisaged to make a cooperative choice in Rapoport (1966, p. 141): "The best outcome for both of us is $(2, 2)$. However if Column assumes that I shall do C, he may well not do C to win the largest payoff. To protect myself I will also refrain from doing C. But this makes for a loss for both of us. Two rational players certainly deserve the outcome $(2, 2)$. I am rational and by the fundamental postulate of game theory, I must assume that Column is also rational. If I have come to the conclusion that C is the rational choice, he too must have come to the same conclusion. Now knowing that he will do C, what shall I do? Shall I not refrain from doing C to get the greatest payoff? But if I have come to this conclusion, he has also probably done so. Again we

[11] The term 'rational' is obviously used here in a different sense from the one used in modern game theory.

end up with $(1,1)$. To ensure that he does not come to the conclusion that he should refrain from C, I better avoid it also. For if I avoid it and am rational, he too will avoid it if he is rational. On the other hand, if rationality prescribes not doing C, then it must also prescribe not doing C for him. At any rate because of the symmetry of the game, rationality must prescribe the same choice to both. But if both choose the same then $(2,2)$ is clearly the better. Therefore, I should choose to do C".

Howard (1988) tests the assertion by Rapoport (1966, p. 139-41) by running a contest between two computer programs. One computer program is designed to play the dominant strategy, defect. Another computer program, called the MIRROR program, is able to recognize if it is playing another MIRROR program, in which case it also cooperates, otherwise it plays defect. There are five copies of each of the programs that play a tournament and, not surprisingly, the MIRROR program achieves higher payoffs. In effect, what the MIRROR program is doing is to replicate the notion that people would cooperate with other like-minded people. In the conclusion, Howard (1988, p. 212) gives an argument that is identical in spirit to the evidential reasoning argument: "If all players use the self-recognition program listed in the Appendix, and play cooperatively only if they recognize their opponents as their twins, then every game will be played cooperatively".

In contrast to the standard explanations, the explanation of cooperation based on evidential reasoning appears to be quite plausible. The fact that a sizeable fraction of the experimental subjects also defect suggests that the results are best accounted for by a mixture in the population of people who use evidential reasoning and causal reasoning. A more formal account of the prisoners' dilemma game under evidential reasoning is give is section 4, below.

2.3 Why do people voluntarily contribute to a public good?

A public good is a *non-rival* and *non-excludable* good. In public good games, a set of individuals simultaneously contribute towards a public good that gives utility to all individuals. In such games the 'first best' that maximizes the joint payoffs is typically achieved when all players contribute. In classical game theory, under causal reasoning, it is optimal for each player to *free ride* (zero contributions). Hence, cooperation cannot be sustained in a Nash equilibrium of the game.

The experimental evidence shows that in early rounds of public good games, participants begin at quite high levels of contributions that are between a half to three-quarters of their maximum possible contributions. Contributions drop off in subsequent rounds unless ex-post punishment of non-cooperators by cooperators is allowed.[12]

Unlike the results of the prisoners' dilemma game, social preferences in the form of

[12] See Dawes and Thaler (1988), Camerer (2003) and Fehr and Gächter (2000).

conditional reciprocity provide a convincing explanation of the experimental results; see, e.g., Fehr and Gächter (2000). In particular, heterogeneity in beliefs provides the leading account of the dynamic decay of cooperation; see, e.g., Fischbacher et al. (2001). In the explanation of the dynamics of such games, the data reveals various degrees of like-minded behaviors of the experimental subjects: (1) 10% always match the expected average contribution of others. (2) 14% fully match the expected average contributions of others until they spend half their endowments. (3) 40% match to a lesser extent. (4) 30% are free riders. (5) 6% have no well defined behavioral pattern. Evidential reasoning may provide the microfoundations of such behaviors. Subjects in the first three categories appear to be using evidential reasoning with various degrees of like mindedness in behaviors. Subjects in category 4 seem to be using causal reasoning. Finally the 6% subjects in the fifth category are not explained by any of these theories. Under evidential reasoning, many players take their own desire to contribute in the first round as diagnostic of the probability with which other players will contribute, hence, they contribute.[13]

Further evidence for evidential reasoning is provided by Gächter and Thöni (2005) who investigate whether cooperation in public goods games with voluntary contribution is higher among *like-minded* people. In order to separate the subjects into like-minded people they initially run a single-round of the public good experiment. The subjects are then grouped by the amount of contributions they made in this round. For instance, the top 3 contributors are grouped into a separate group (the TOP group) as having the greatest inclination to contribute. The public good game is then played separately in each group. Over the next 10 rounds, contributions are much higher and free riding much lower in the TOP group, which achieves nearly the first best level in several rounds.[14] Evidential reasoning would predict that when high contributors have greater confidence that they are grouped with like-minded people, they contribute more.

2.4 Evidential reasoning about states of other humans

In an experiment conducted by Van Boven and Loewenstein (2005), participants were asked to imagine a hiker who is lost in the woods and asked whether the hiker was more likely to be thirsty or hungry. When the experimental participants were made to experience these states (thirst or hunger) they attributed similar states to the hypothetical hiker despite their own states having no correlation or cue value with the state of the lost hiker. In each case, they assign *diagnostic significance* to their own state in inferring the state of the lost hiker.

[13] For the dynamic version of the public good game we conjecture that evidential reasoning in conjunction with negative reciprocity gives a better description of the evidence from public good games.

[14] Even the endgame effect, i.e., the sharp drop in contributions in the last experiment is most pronounced among the bottom group.

Lenton et al. (2007) described a common scenario to a group of experimental subjects. The scenario involved two students (one male and the other female) dining out on a date and then going to the woman's apartment and listening to music. The subjects were then asked to assign a probability to the event that the two people in the scenario would have casual sex. Those who were personally more predisposed to casual sex also assigned a higher probability to the event. People seem to assign diagnostic significance to their own actions in inferring the actions of others although they know that their actions do not cause others to take specific actions.

Evidential reasoning may also aid in the emergence of empathy. When I observe a person with a broken arm, I might use evidential reasoning to infer that the other person suffers the same pain that I did when I had a broken arm.

2.5 Why do people vote and how do they form beliefs?

Under causal reasoning, if voters do not derive utility purely from the act of voting, then, given that any one voter is most unlikely to be pivotal, nobody should vote. But then why do so many people vote? This is the *voting paradox*. The situation faced by a voter who uses causal reasoning is depicted in Table 3 (see, for instance, Grafstein, 1991).

	D	R
A	pv_D	$(1-p)v_D$
V	$pv_D - c$	$(1-p)v_D - c$

Table 3: Voting game

In Table 3 there are two political parties, Democrats (D) and Republicans (R). In a Presidential election, a typical voter has an action set $\{A, V\}$ where A is the action 'abstain from voting' and V is the action 'vote'. Voting entails a fixed cost $c > 0$. The voter assigns a probability $p \geq 0$ that the Democrat candidate wins and the complementary probability that the Republican candidate wins.[15] The voter gets utility $v_D > 0$ from a Democrat candidate and $v_R > 0$ from a Republican candidate. Clearly for any p, c, v_D, v_R it is a dominant strategy to abstain from voting given that no voter is pivotal.

The explanation of voting when a voter uses evidential reasoning is as follows. "If I do not vote for my preferred party, then probably like-minded people will not vote, and my preferred party will lose to the other party. On the other hand, if I decide to vote then, probably, other like-minded people will also make a similar decision and my party

[15] Since information on Presidential elections is publicly and equally available to all voters, one may even make the strong argument that all voters assign the same probability, p, but this is not critical to the argument below.

Year	Presidential Candidates	% of Democrat voters expecting Democrat win	% of Republican voters expecting Republican win
1988	Dukakis vs. Bush	51.7	94.2
1984	Mondale vs. Reagan	28.8	99.9
1980	Carter vs. Reagan	87.0	80.4
1976	Carter vs. Ford	84.2	80.4
1972	McGovern vs. Nixon	24.7	99.6
1968	Humphrey vs. Nixon	62.5	95.4
1964	Johnson vs. Goldwater	98.6	30.5
1960	Kennedy vs. Nixon	78.4	84.2
1956	Stevenson vs. Eisenhower	54.6	97.6
1952	Stevenson vs. Eisenhower	81.4	85.9

Table 4: Source: Table 6 in Forsythe et al. 1992

has a better chance of winning. So I vote if I wish my party to win, otherwise I do not."[16] In each case, the binary voting decision (vote or not vote) has diagnostic significance in forming beliefs about whether others will vote, although it is critical to note that *one's action to vote does not cause others to vote.*

As Krueger and Acevedo (2008, p. 468) put it: "Compared with a Republican who abstains, for example, a Republican who votes can be more confident that other Republicans vote in large numbers". Quattrone and Tversky (1984), Grafstein (1991) and Koudenburg et al. (2011) show that experimental evidence is strongly supportive of this view. Delavande and Manski (2012) argue that state and national poll information in the US is readily available public knowledge. On the other hand, private knowledge in elections is likely to be very limited. Hence, all individuals should form identical estimates of the winning probabilities of the political parties. In contrast to these expectations, they find strong support for evidential reasoning. Voters who vote, assign too high a probability to their preferred party of winning the elections. Their findings are invariant with respect to males/females, whites/non-whites, educated/non-educated etc. Hence, and consistent with earlier work, there is a strong possibility that evidential reasoning is hard-wired in humans.

Consider Table 4 that reports survey evidence from successive US Presidential elections that is supportive of the evidential reasoning explanation. Voters who intend to vote Democrats typically assign high probabilities to the Democrat candidate winning. By contrast, voters who intend to vote Republican assign high probabilities to a Republican win. Thus, voters seem to take their own actions as diagnostic of what other like-minded

[16] Other possible explanations for voting, for instance, that people vote out of a sense of civic duty cannot explain several kinds of strategic voting and the variation in voter turnout when an election is believed to be close; see Krueger and Acevedo (2008).

people will do. These findings show that, in some circumstances at least, people behave *as if* their actions were causal even when they are merely diagnostic or evidential.

2.6 Coordinated attack: An application to the battle of Waterloo

Consider the following version of the well known historical *coordinated attack* problem; see for instance, Halpern (1986). Wellington (W) and Blucher (B) wish to attack their common enemy Napoleon (N). If W or B attack on their own, N will win. But if W and B attack together, they will win. W sends a message to B saying he will attack, but only if he receives confirmation from B that B will also attack. B replies that he will attack, but only if he receives confirmation that his message has reached W, and so on. Under causal reasoning, neither W nor B will attack. The situation is shown in Table 5. W is the row player and B is the column player. The payoffs when both players simultaneously Attack are v, v, $v > 0$. If one of the generals choose Attack and the other chooses Capitulate then the respective payoffs are $-C, -c$, $C > c > 0$.

	Attack	Capitulate
Attack	v, v	$-C, -c$
Capitulate	$-c, -C$	$-c, -c$

Table 5: The coordinated attack problem

The game has two pure strategy Nash equilibria: (Attack, Attack) and (Capitulate, Capitulate). There is also a mixed strategy Nash equilibrium in which each player plays Attack with probability $\frac{C-c}{v+C}$. Under causal reasoning, each player finds it profitable to Attack if he believes that the other will Attack with probability $p \geq \frac{C-c}{v+C}$. In general there is no mechanism that enables players to coordinate under causal reasoning.

Under evidential reasoning, W and B may both choose Attack if they perceive each other to be like-minded. Both might reason as follows: "The other player is like-minded so he must realize that if we both choose Capitulate then we will get a lower payoff relative to the case where we both choose Attack. So the other player is very likely to choose Attack. Hence, I also choose Attack". Each player may use his own reasoning as diagnostic evidence that the other will choose Attack (and, maybe, a finite number of messages is sufficient to enforce this psychological mode of reasoning). Eventually, Wellington and Blucher did attack Napoleon with decisive consequences in the Battle of Waterloo. Thus evidential reasoning, while not a strictly correct method of reasoning may, nevertheless, be a useful heuristic that has practical utility in coordinating actions in static games.

3 Evidential equilibrium in static games of complete information

We now consider static games of complete information. Section 6 considers static games of incomplete information.

3.1 Elements of standard game theory

Consider the following standard description of a static game of complete information, $\{N, \mathbf{A}, \boldsymbol{\pi}\}$. $N = \{1, 2, ..., n\}$ is the set of *players*. $A_i \subseteq \mathbb{R}$ is the set of *actions* open to player i. We denote a typical member of A_i by a_i.[17] $\mathbf{A} = \times_{i=1}^{n} A_i$ gives all possible action profiles of the players. $\mathbf{A}_{-i} \subseteq \mathbb{R}^{n-1}$ is the set of vectors of actions open to the other players. Denote by Δ_i, the set of probability distributions over the set of actions A_i. We denote a typical element of Δ_i by σ_i and call it a *strategy*. $\sigma_i(a_i)$ is the probability with which player i plays $a_i \in A_i$, so $\sigma_i(a_i) \geq 0$ and $\sum_{a_i \in A_i} \sigma_i(a_i) = 1$. In particular, if $\sigma_i(a_i) = 1$ (hence, $\sigma_i(a_i') = 0$ for $a_i' \neq a_i$), then we call σ a *pure strategy* and we identify it with the action a_i.

A profile of strategies of all players is denoted by $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, ..., \sigma_n) \in \boldsymbol{\Delta}$, where $\boldsymbol{\Delta} = \times_{i=1}^{n} \Delta_i$ is the set of all possible profiles of strategies. A particular profile of strategies of other players is denoted by $\boldsymbol{\sigma}_{-i} = (\sigma_1, ..., \sigma_{i-1}, \sigma_{i+1}, ..., \sigma_n) \in \boldsymbol{\Delta}_{-i} = \times_{j \in N - \{i\}} \Delta_j$. The payoff of player i is $\pi_i : \boldsymbol{\Delta} \to \mathbb{R}$ and $\boldsymbol{\pi}$ is the vector of *payoffs*. Given a strategy profile, $\boldsymbol{\sigma} = (\sigma_i, \boldsymbol{\sigma}_{-i}) \in \boldsymbol{\Delta}$, the payoff to player i is $\pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) \in \mathbb{R}$. The structure of the game, $\{N, \mathbf{A}, \boldsymbol{\pi}\}$, is common knowledge among the players. In an experimental setup, common knowledge can be achieved by a public announcement of $\{N, \mathbf{A}, \boldsymbol{\pi}\}$. This is the sense in which this is a game of complete information. However, when each player, i, chooses his strategy, σ_i, he does not know the strategies, $\boldsymbol{\sigma}_{-i}$, chosen simultaneously by the other players and this is the sense in which this is a static simultaneous move game.

Definition 1 *(Nash, 1951): A strategy profile $\boldsymbol{\sigma}^* = (\sigma_1^*, \sigma_2^*, ..., \sigma_n^*) \in \boldsymbol{\Delta}$ is a Nash equilibrium in the game, $\boldsymbol{\Gamma} = \{N, \mathbf{A}, \boldsymbol{\pi}\}$, if σ_i^* maximizes $\pi_i\left(\sigma_i, \boldsymbol{\sigma}_{-i}^*\right)$ with respect to σ_i, given $\boldsymbol{\sigma}_{-i}^*$, for each $i \in N$, i.e.,*

$$\pi_i\left(\sigma_i^*, \boldsymbol{\sigma}_{-i}^*\right) \geq \pi_i\left(\sigma_i, \boldsymbol{\sigma}_{-i}^*\right) \text{ for all } \sigma_i \in \Delta_i.$$

Note that there is no role for beliefs about the strategies of others in the game $\{N, \mathbf{A}, \boldsymbol{\pi}\}$, nor in the definition of a Nash equilibrium (Definition 1). Hence, we augment the game, $\{N, \mathbf{A}, \boldsymbol{\pi}\}$, with a profile of "social projection functions", $\mathbf{P}$ that specify the beliefs of players; this is undertaken in subsection 3.2, below.

[17] For this paper it will suffice to take an action for a player to be a real number. More generally, an action may be a vector of real numbers or an even more abstract entity.

3.2 Social projection functions

We would like to define a function that captures the beliefs that a player has about the strategies of the other players, conditional on his own strategy. We will call such a function a *social projection function*.[18]

Definition 2 *(Social projection functions, SPF): A social projection function for player i (SPF for short), is a mapping $\mathbf{P}_i : \Delta_i \to \mathbf{\Delta}_{-i}$, that assigns to each strategy, $\sigma_i \in \Delta_i$, for player i, an $(n-1)$ vector of strategies for the other players. We write $P_{ij}(a_j|\sigma_i)$ for the 'subjective belief' that player i assigns to player j playing $a_j \in A_j$, conditional on player i playing $\sigma_i \in \Delta_i$. Hence, $P_{ij}(a_j|\sigma_i) \geq 0$ and $\sum_{a_j \in A_j} P_{ij}(a_j|\sigma_i) = 1$. We may write $\mathbf{P}_i(\sigma_i) = \boldsymbol{\sigma}_{-i}^e(\sigma_i)$ to indicate that $\mathbf{P}_i(\sigma_i)$ is the $(n-1)$ vector of strategies that player i anticipates that the other players will follow if player i adopts the strategy σ_i.*

We now define *causal reasoning*, the dominant type of reasoning used in game theory, and *evidential reasoning*.

Definition 3 *(Causal Reasoning): We say that player i uses causal reasoning if $P_{ij}(a_j|\sigma_i)$ is independent of σ_i for each $a_j \in A_j$ and each $j \neq i$, i.e., if $\mathbf{P}_i(\sigma_i) = \mathbf{P}_i(\sigma_i')$ for all $\sigma_i, \sigma_i' \in \Delta_i$.*

Definition 4 *(Evidential Reasoning) We say that player i uses evidential reasoning if it is not necessarily the case that $\mathbf{P}_i(\sigma_i) = \mathbf{P}_i(\sigma_i')$ for all $\sigma_i, \sigma_i' \in \Delta_i$.*

We make some important observations about Definitions 2-4 in the next remark.

Remark 1 *: (a) In a static game of full information, players are uncertain of the actions taken by others. Under evidential reasoning, player i resolves this uncertainty by assigning diagnostic significance to his own choice of strategy, σ_i, in inferring the strategies of the other players, $\boldsymbol{\sigma}_{-i}$, using his social projection function, $\mathbf{P}_i$. For this reason, Definition 4 allows for $\mathbf{P}_i(.|\sigma_i)$ to change as σ_i changes. However, it is important to realize that there is no causal connection between σ_i and $\boldsymbol{\sigma}_{-i}$. The choice of σ_i by player i merely influences his 'belief' about the strategies, $\boldsymbol{\sigma}_{-i}$, of the other players. In particular, players who use evidential reasoning know that their own actions have no causal effects in altering the actions of others when they change their own actions. Further, a SPF (Definition 2) specifies the beliefs of a player for all possible actions of others, including out-of-equilibrium actions. The beliefs of a player need not turn out to be fulfilled in equilibrium just as in*

[18] We use the term *social projection function* because, on the one hand, it is obviously connected with social projection and evidential reasoning and, on the other hand, to distinguish it from the term *projection function* as commonly used in mathematics.

other disequilibrium-in-beliefs models such as the Level-k model.
(b) If player i uses causal reasoning as in classical game theory (see Definition 3) then he assigns no diagnostic significance to his own strategy, σ_i, in inferring the strategies, $\boldsymbol{\sigma}_{-i}$, followed by the other players.[19] *Thus, under causal reasoning, $\mathbf{P}_i(\sigma_i)$ remains fixed as σ_i changes. From Definitions 3, 4, causal reasoning is a special case of evidential reasoning*

Players need not impute diagnostic significance to their actions when others are perceived not to be like-minded; the next definition formalizes this idea. This issue is considered further in subsection 3.3.

Definition 5 *(Ingroups and Outgroups): Suppose that players use evidential reasoning.*
(a) Player i regards player j $(j \neq i)$ as an outgroup member if $P_{ij}(a_j|\sigma_i)$ is independent of σ_i, i.e., if $P_{ij}(a_j|\sigma_i) = P_{ij}(a_j|\sigma_i')$ for all $\sigma_i, \sigma_i' \in \Delta_i$ and all $a_j \in A_j$. Otherwise, player i regards player j $(j \neq i)$ as an ingroup member.
(b) Let $M \subset N$ be a non-empty set of players. If every player in M regards every other player in M as an ingroup member, then M is an ingroup.
(c) Let $L \subset N$ and $M \subset N$ be disjoint non-empty sets of players. Suppose every player in L regards every player in M as an outgroup member. Then we say that M is an outgroup relative to L.

Remark 2 *: Since player i plays action $a_i \in A_i$ with probability $\sigma_i(a_i)$ and believes that player j will play action $a_j \in A_j$ with probability $P_{ij}(a_j|\sigma_i)$ (the latter is conditional on σ_i), it follows that player i also believes that the joint probability of a_i and a_j being played is $P_{ij}(a_i, a_j|\sigma_i) = \sigma_i(a_i) P_{ij}(a_j|\sigma_i)$. Suppose that player i regards player j as an outgroup member. Then (and only then) $P_{ij}(a_j|\sigma_i)$ is independent of $\sigma_i \in \Delta_i$. In this case we can set $P_{ij}(a_j|\sigma_i) = \sigma_j^e(a_j)$ (which, of course, depends on i but is independent of σ_i). Hence, in this case, $P_{ij}(a_i, a_j|\sigma_i) = \sigma_i(a_i) P_{ij}(a_j|\sigma_i) = \sigma_i(a_i)\sigma_j^e(a_j)$. Thus, if player i regards player j as an outgroup member, then player i believes that the probability with which he (player i) plays $a_i \in A_i$ is independent from the probability that he believes j will play $a_j \in A_j$. In particular, if player i uses causal reasoning, then he regards all others as outgroup members and, hence, he believes that his actions are independent of the actions of all other players.*

Many different types of social projection functions (SPF's) are possible. There are possibly various degrees of like-mindedness. However, the examples in the introduction and in Section 2 above, suggest a particularly salient SPF that imparts the main predictive

[19] By contrast, in a dynamic game (under causal reasoning), if (say) player 1 moves first, choosing the strategy σ_1, followed by player 2 who chooses strategy σ_2, having observed a realization of σ_1, then σ_2 may very well depend on σ_1. When choosing σ_1, player 1 will take into account the influence of his choice on the future behaviour of player 2. This should not be confused with evidential reasoning.

content to theory. This SPF seems important when choices are low dimensional such as the binary choices 'cooperate or defect in a prisoners' dilemma game' or 'vote Democrats or Republicans in US Presidential elections' or 'coordinate or fail to coordinate in coordination games' etc. In each of these cases, players may think that other like-minded players will choose *identical actions* to their own. The corresponding SPF in these cases is the *identity social projection function*. In this case, conditional on playing any strategy σ_i, player i believes that like-minded players will play an identical strategy.

Definition 6 *(Identity social projection function): Let $M \subseteq N$ be a subset of players. Suppose that all players in M have the same action set, i.e., $A_i = A_j = A$ for all $i, j \in M$. Let $\mathbf{P}_i$ be the social projection function for player $i \in M$. Recall that $P_{ij}(a|\sigma_i)$ is the probability that player i assigns to player j playing action a when the strategy of player i is given by σ_i. If $P_{ij}(a|\sigma_i) = \sigma_i(a)$ for all $a \in A$ and all $j \in M - \{i\}$, then we say that $\mathbf{P}_i$ is an identity social projection function on M. If $M = N$, then we say that $\mathbf{P}_i$ is an identity social projection function.*

Definition 7 *(Perfect ingroups): Let $M \subseteq N$ be a subset of players. Suppose that all players in M have the same action set, i.e., $A_i = A_j = A$ for all $i, j \in M$. Let $\mathbf{P}_i$ be the social projection function for player $i \in M$. If $\mathbf{P}_i$ is an identity social projection function on M, for each player $i \in M$, then M is a perfect ingroup.*

Definition 8 *(Evidential game): Consider the static game of complete information, $\{N, \mathbf{A}, \boldsymbol{\pi}\}$. Let $\mathbf{P} = (\mathbf{P}_1, \mathbf{P}_2, ..., \mathbf{P}_n)$ be a profile of social projection functions, where $\mathbf{P}_i$ is the social projection function of player $i \in N$ (Definition 2). Then we denote the game augmented with the vector of social projection functions, $\mathbf{P}$, by $\Gamma = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$ and we call it an evidential game. We say that players in such a game use evidential reasoning. In particular, if each $\mathbf{P}_i(\sigma_i)$ is independent of σ_i, then we say that Γ is a causal game.*

Remark 3 *: (a) From Definition 8 a causal game is a special case of an evidential game. (b) Suppose $\mathbf{P}_i(\sigma_i)$ is independent of σ_i, for each player, i, so that $\Gamma = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$ is a causal game. Γ is still richer than the static game of complete information, $\{N, \mathbf{A}, \boldsymbol{\pi}\}$, because Γ incorporates players' beliefs about other players' actions, as given by $\mathbf{P}$.*

Example 4 *: Consider the matching pennies game.*

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

The set of players is $N = \{1, 2\}$. The action sets are $A_1 = A_2 = \{H, T\}$. Player

1, the row player, plays H and T with respective probabilities p, $1-p$. Player 2, the column player, plays H and T with respective probabilities q, $1-q$. The sets of possible strategies are $\Delta_1 = \{p : 0 \leq p \leq 1\}$ for player 1 and $\Delta_2 = \{q : 0 \leq q \leq 1\}$ for player 2. For any profile of strategies (p, q), $p, q \in [0, 1]$, the payoff functions of the players are $\pi_1(p, q) = -(1-2p)(1-2q)$ and $\pi_2(p, q) = (1-2p)(1-2q)$. The following are examples of social projection functions

$$\begin{cases} P_{12}(H \mid p) = p,\ P_{12}(T \mid p) = 1-p \text{ for all } p \in [0,1], \\ P_{21}(H \mid q) = 0.5,\ P_{21}(T \mid q) = 0.5 \text{ for all } q \in [0,1]. \end{cases} \tag{2}$$

According to (2), player 1, who uses evidential reasoning, believes that if he (player 1) plays H with probability p, then so will player 2 for any $p \in [0, 1]$. Hence, player 1 has an identity social projection function. It is critical to note that these are the 'beliefs' of player 1. There is no presumption that these beliefs will turn out to be justified ex-post. Player 2, who uses causal reasoning, believes that player 1 will play H with probability 0.5, whatever strategy, q, player 2 chooses. Hence, player 1 regards player 2 as an ingroup member but player 2 regards player 1 as an outgroup member. Hence, $N = \{1, 2\}$ fails to be an ingroup. On the other hand, if both players had identity social projection functions, then $N = \{1, 2\}$ would be an ingroup (in fact, a perfect ingroup). By contrast, if in (2) we had, say, $P_{12}(H \mid p) = 0.3$ for all $p \in [0, 1]$, then both players would exhibit causal reasoning; and this example would become a causal game as in classical game theory.

3.3 Ingroups, outgroups and evidential reasoning

Intuitively, an "ingroup" is a group of players each of whom believes that the others are like-minded and, hence, would behave in a similar, but not necessarily identical, manner. The literature has typically assumed that players do not use their actions as diagnostic of the actions for "outgroup" players; see, for instance, Krueger (2007), Robbins and Krueger (2005); and our definitions reflect this (see, in particular, Definition 5).

However, recent evidence suggests a more nuanced view that is also consistent with our definitions. Koudenburg et al. (2011) show that voters project their own preference for a political party to non-voters even when they are informed about the poll results for non-voters. Thus, voters may regard non-voters as ingroup members, though in a strict sense only the set of voters may be thought to form an ingroup (Definition 5).

Riketta and Sacramento (2008) cite several references to show that members of an ingroup assign beliefs about other members even when they could have no possible information about those members (recall subsection 2.4). They find that an ingroup member may have a harmonious (or cooperative) relation with other members. On the other hand, they also find that an ingroup member may be in competition (or conflict) with other members. In the latter case, an ingroup member may believe that the actions of others are in contrast[20] to his own actions (the *contrast effect*). Subsection 5.2 will give further examples of these.

3.4 Equilibria

In this section we consider the relevant solution concept for static evidential games of complete information that we call an evidential equilibrium.

Definition 9 *(Optimal strategies): An optimal strategy for player i, $\sigma_i^* \in \Delta_i$, in the evidential game $\Gamma = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$ (Definition 8), is one that maximizes the payoff function, $\pi_i(\sigma_i, \mathbf{P}_i(\sigma_i))$ (i.e., $\pi_i\left(\sigma_i, \boldsymbol{\sigma}_{-i}^e(\sigma_i)\right)$) of player i.*

Definition 10 *(Evidential equilibria): The strategy profile $\boldsymbol{\sigma}^* = (\sigma_1^*, \sigma_2^*, ..., \sigma_n^*) \in \boldsymbol{\Delta}$ is an evidential equilibrium of the evidential game $\boldsymbol{\Gamma} = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$ if σ_i^* is an optimal strategy for each $i \in N$ (Definition 9).*

Definitions 9, 10 identify an important feature of an evidential equilibrium. In static games of complete information, when, for whatever reason, there is uncertainty about what others will do, evidential reasoning converts an essentially strategic situation to a decision-theoretic problem. It does this through the incorporation of a social projection function as an essential part of the game. This appears to be consistent with the evidence (see section 2).

The optimal actions of each player in an evidential equilibrium are found from solving a decision-theoretic problem in what is classically a strategic game. There are no higher order beliefs. Each player does not think about strategically exploiting the SPF of the other player. Indeed the game $\boldsymbol{\Gamma} = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$ does not involve any assumptions about the mutual or common knowledge of $\mathbf{P}$. Requiring π_i to be a continuous function on the compact mixed strategy space guarantees that an equilibrium exists in the decision theoretic problem.

Note that Definition 10 only requires that a strategy for a player be optimal given *his beliefs*. But, of course, beliefs may not turn out to be correct, ex-post. Ultimately, the choice among models in all science is guided by the evidence. The evidence reviewed above (and below) shows that in static games, beliefs about others often turn out to be incorrect.[21] Nevertheless, it is of interest to consider the special case where beliefs turn out to be correct, at least in equilibrium. This is the subject of the next two definitions.

[20] In the case of binary actions, A, B, if a player wishes to play A then the action that is in contrast to his own is B.

[21] Even in experiments, where successive rounds of play lead to an improvement in the accuracy of beliefs, one may be interested in explaining the behaivor in early rounds of play where beliefs do not turn out to be correct. Often such behavior mimics real life situations in which decision makers do not get repeated or frequent opportunities to make their decisions.

Definition 11 *(Mutually consistent strategies): A strategy profile* $\boldsymbol{\sigma}^* = (\sigma_1^*, \sigma_2^*, ..., \sigma_n^*) \in \boldsymbol{\Delta}$ *of the evidential game* $\boldsymbol{\Gamma} = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$ *(Definition 8) is a mutually consistent vector of strategies if* $\mathbf{P}_i(\sigma_i^*) = \boldsymbol{\sigma}_{-i}^*$ *for all* $i \in N$*, i.e., if* $P_{ij}(a_j|\sigma_i^*) = \sigma_j^*(a_j)$*, for all* $i, j \in N$*,* $i \neq j$*, and all* $a_j \in A_j$*.*

In other words, a strategy profile $\boldsymbol{\sigma}^* = (\sigma_1^*, \sigma_2^*, ..., \sigma_n^*)$ is a mutually consistent vector of strategies, if for all players $i, j \in N$, $i \neq j$ and all actions, a_j, open to player j, the probability $P_{ij}(a_j|\sigma_i^*)$ that player i 'believes' that player j will play action a_j (given σ_i^*) is equal to the 'actual' probability $\sigma_j^*(a_j)$ with which player j plays a_j. Indeed, imposing such a social projection is quite strong, as are the epistemic conditions for a Nash equilibrium; the relation between the two is shown in Section 3.5 below. The empirical evidence (see Section 2 above) does not support the imposition of such social projection functions.

Definition 12 *(Consistent evidential equilibria): A consistent evidential equilibrium of the evidential game* $\boldsymbol{\Gamma} = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$ *is an evidential equilibrium,* $\boldsymbol{\sigma}^* \in \boldsymbol{\Delta}$*, which is also a mutually consistent vector of strategies (Definitions 10 and 11).*

3.5 Nash equilibria and consistent evidential equilibria

As one might expect, there is a natural correspondence between Nash equilibria and consistent evidential equilibria. This is formally stated and established by the following proposition, which is a special case of the famous result of Aumann and Brandenburger (1995) on the epistemic foundations of a Nash equilibrium.

Proposition 1 *: (a) Let* $\boldsymbol{\sigma}^* \in \boldsymbol{\Delta}$ *be a Nash equilibrium in the static game of complete information,* $\{N, \mathbf{A}, \boldsymbol{\pi}\}$*. Consider the (constant) social projection functions:* $\mathbf{P}_i(\sigma_i) = \boldsymbol{\sigma}_{-i}^*$*,* $i \in N$*. Then* $\boldsymbol{\sigma}^*$*, is a consistent evidential equilibrium in the evidential game* $\Gamma = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$*. Furthermore,* Γ *is a causal game.*
(b) Let $\boldsymbol{\sigma}^* \in \boldsymbol{\Delta}$ *be an evidential equilibrium in the evidential game* $\Gamma = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$*, where* $\mathbf{P}$ *is the profile of constant social projection functions* $\mathbf{P}_i(\sigma_i) = \boldsymbol{\sigma}_{-i}^*$*,* $i \in N$ *(hence,* $\boldsymbol{\sigma}^*$ *is a consistent evidential equilibrium and* Γ *is a causal game). Then* $\boldsymbol{\sigma}^*$ *is a Nash equilibrium in the static game of complete information* $\{N, \mathbf{A}, \boldsymbol{\pi}\}$*.*

Proof of Proposition 1: (a) Let $\boldsymbol{\sigma}^* \in \boldsymbol{\Delta}$ be a Nash equilibrium in the static game of complete information, $\{N, \mathbf{A}, \boldsymbol{\pi}\}$. Consider the social projection functions: $\mathbf{P}_i(\sigma_i) = \boldsymbol{\sigma}_{-i}^*$, $i \in N$. Since $\boldsymbol{\sigma}^*$ is a Nash equilibrium (Definition 1), it follows that σ_i^* maximizes $\pi_i\left(\sigma_i, \boldsymbol{\sigma}_{-i}^*\right)$ with respect to σ_i, given $\boldsymbol{\sigma}_{-i}^*$, for each $i \in N$. Since, by construction, $\mathbf{P}_i(.|\sigma_i) = \boldsymbol{\sigma}_{-i}^*$, $i \in N$, it follows that σ_i^* maximizes $\pi_i(\sigma_i, \mathbf{P}_i(\sigma_i))$ with respect to σ_i, for each $i \in N$. Hence, $\boldsymbol{\sigma}^*$ is an evidential equilibrium (Definitions 9 and 10) in the evidential game $\Gamma = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$. Furthermore, since, by construction, $\mathbf{P}_i(\sigma_i) = \boldsymbol{\sigma}_{-i}^*$, $i \in N$, it follows that $\boldsymbol{\sigma}^*$ is a consistent evidential equilibrium (Definitions 11 and 12). Since $\mathbf{P}$ is a profile of constant social projection functions, it follows that $\Gamma = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$ is a causal game (Definition 8).

(b) Let $\boldsymbol{\sigma}^* \in \boldsymbol{\Delta}$ be an evidential equilibrium in the evidential game $\Gamma = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$, where $\mathbf{P}$ is the profile of constant social projection functions $\mathbf{P}_i(\sigma_i) = \boldsymbol{\sigma}^*_{-i}$, $i \in N$. Then σ_i^* maximizes $\pi_i(\sigma_i, \mathbf{P}_i(\sigma_i))$ with respect to σ_i, for each $i \in N$ (Definitions 9 and 10). But $\mathbf{P}_i(\sigma_i) = \boldsymbol{\sigma}^*_{-i}$, $i \in N$, hence σ_i^* maximizes $\pi_i\left(\sigma_i, \boldsymbol{\sigma}^*_{-i}\right)$ with respect to σ_i, for each $i \in N$. Hence, $\boldsymbol{\sigma}^*$ is a Nash equilibrium in the static game of complete information $\{N, \mathbf{A}, \boldsymbol{\pi}\}$ (Definition 1). $\blacksquare$.

4 The prisoners' dilemma game

We now use the prisoners' dilemma game in Table 6 to illustrate some of the key concepts developed so far. We call the row player as player 1 and the column player as player 2. We consider the following cases.

1. Both players use evidential reasoning.
2. One player uses evidential reasoning but the other uses causal reasoning.
3. Both players use causal reasoning but beliefs turn out to be wrong ex-post.
4. Both players use causal reasoning and beliefs turn out to be correct ex-post.

The last case illustrates Proposition 1a, namely, that there will always be a profile of social projection functions for which a given Nash equilibrium corresponds to a consistent evidential equilibrium. A comparison of Cases 3 and 4 shows that causal reasoning does not guarantee mutual consistency of beliefs.

	C	D
C	$2,2$	$0,3$
D	$3,0$	$1,1$

Table 6: The Prisoner's Dilemma Game

Here $N = \{1,2\}$, $A_1 = A_2 = \{C, D\}$, $\mathbf{A} = \{C, D\} \times \{C, D\}$ and $\boldsymbol{\pi}$ is given by the payoff matrix in Table 6. Each player has a dominant action, D, thus, the unique Nash equilibrium of this game is (D, D) (Definition 1). By contrast, the empirical evidence reviewed in subsection 2.2, above, shows that 50% or more of the outcomes involve the play (C, C). We can set up this game as either an evidential game or a causal game. Since the Nash equilibrium for a prisoners' dilemma game is in pure strategies, we focus only on pure strategies.

From Definition 4, in general, $P_{ij}(.|\sigma_i)$ will vary with σ_i but, if player i uses causal reasoning, as in classical game theory, then $P_{ij}(.|\sigma_i)$ will be independent of σ_i (Definition 3).

Case 1: In this case the SPF for each player is given by

$$\begin{cases} P_{12}(C|C) = 1, P_{12}(D|C) = 0; P_{12}(D|D) = 1, P_{12}(C|D) = 0, \\ P_{21}(C|C) = 1, P_{21}(D|C) = 0; P_{21}(D|D) = 1, P_{21}(C|D) = 0. \end{cases} \tag{3}$$

From (3), both players use evidential reasoning, so this is an evidential game. In particular, each player uses his *identity social projection function* (Definition 6). Together, both players form an ingroup. Given the social projection function for player 1, his unique optimal choice is C. Similarly, C is also the optimal choice for player 2. Hence, (C, C) is an evidential equilibrium (Definition 10). Furthermore, (C, C) is the unique evidential equilibrium of this game. Each player expects the other to play C in response to C, which turns out to be correct, ex-post. Therefore, (C, C) is a mutually consistent vector of strategies (Definition 11). Hence, (C, C) is a consistent evidential equilibrium (Definition 12). In contrast, (C, C) is not the Nash equilibrium of the game. Indeed, (C, C) requires each player to play a strictly dominated strategies. However, (C, C) is Pareto optimal. Note that under evidential reasoning one does not need repeated game arguments to justify cooperation in the static prisoners' dilemma game. Moreover, this is consistent with the play of more than 50% of players (see Section 2.2 above).

Readers trained in classical game theory may wish to make the following argument. Since $P_{12}(C|C) = 1$ and players cannot change the actions of opponents by changing their action, why does not player 1 defect and increase his outcome to 3, hence, breaking the cooperative equilibrium? The answer to this is given in section 2.2.2, above; we give a brief summary here. It is also the case that $P_{12}(D|D) = 1$. Player 1 knows that player 2, who is like-minded will also benefit by thinking in an analogous manner and, hence, both will defect, in which case both get a payoff of 1. However, this does not maximize the payoffs, which is required for an evidential equilibrium (see Definitions 9, 10). Hence, player 1 finds it worthwhile to cooperate. For analogous reasons, player 2 also cooperates. Thus, mutual cooperation is sustained.

Remark 4 : *We now make a set of observations that apply to each of the four cases that we consider in this subsection. The SPF for each player in (3) also specifies out-of-equilibrium beliefs. For instance, $P_{12}(D|C) = 0$ specifies the belief of player 1 that if he were to find it optimal to play C then the probability that player 2 will play D is zero. It does not mean that by playing C, player 1 can induce player 2 to play D; indeed there is no such causal link. The optimal actions of each player, (C, C) in this case, are found from solving a decision-theoretic problem, using the heuristic of evidential reasoning, and none of the players attempt to strategically exploiting the SPF of the other player. Indeed* *there is no requirement in an evidential game that there even be mutual knowledge of the SPF.*

Case 2. In this case, the SPF for each player is given by

$$\begin{cases} P_{12}(C|C) = 1, P_{12}(D|C) = 0; P_{12}(D|D) = 1, P_{12}(C|D) = 0, \\ P_{21}(C|C) = 1, P_{21}(D|C) = 0; P_{21}(C|D) = 1, P_{21}(D|D) = 0. \end{cases}$$

Player 1 uses evidential reasoning and, in particular, his identity social projection function, as in case 1 above. Player 2, on the other hand, uses causal reasoning and, in particular, mistakenly assumes that player 1 will always cooperate. This is an evidential game. The unique evidential equilibrium (Definition 10) is (C, D). It is an evidential equilibrium because each player's chosen action is optimal, given his beliefs, which are captured by his social projection function. It is not a consistent evidential equilibrium because the belief of player 1 turns out to be mistaken in equilibrium ($P_{12}(C|C) = 1$ but player 2 plays D instead). By contrast, the belief of player 2 that player 1 always plays C turns out to be correct in equilibrium.

Case 3. In this case the SPF for each player is given by

$$\begin{cases} P_{12}(C|C) = 0, P_{12}(D|C) = 1; P_{12}(D|D) = 0, P_{12}(C|D) = 1, \\ P_{21}(C|C) = 0, P_{21}(D|C) = 1; P_{21}(C|D) = 1, P_{21}(D|D) = 0. \end{cases}$$

Given these social projection functions, the unique payoff maximizing strategy for each player is to play D (Definition 9). Hence, (D, D), is the unique evidential equilibrium (Definition 10). However, (D, D), is not a mutually consistent vector of strategies (Definition 11) because each player expects his opponent to play C in response to D but the opponent's response is D. Hence (D, D) is not a consistent evidential equilibrium (Definition 12). The importance of this case arises because it is sometimes believed that a Nash equilibrium under causal reasoning requires mutual consistency of beliefs.

Case 4. In this case, the SPF for each player is given by

$$\begin{cases} P_{12}(C|C) = 0, P_{12}(D|C) = 1; P_{12}(D|D) = 1, P_{12}(C|D) = 0, \\ P_{21}(C|C) = 0, P_{21}(D|C) = 1; P_{21}(C|D) = 0, P_{21}(D|D) = 1. \end{cases}$$

Both players use causal reasoning, so this is a *causal game*. Given his social projection function, playing D is the unique optimal strategy for player 1 (Definition 9). And similarly for player 2. Hence (D, D) is the unique evidential equilibrium (Definition 10). Furthermore, (D, D) is a mutually consistent vector of strategies (Definition 11) because each player expects his rival to play D and, in fact, his rival does play D. Hence, (D, D) is a consistent evidential equilibrium (Definition 12). The unique Nash equilibrium of this game is, of course, (D, D). Hence this case illustrates Proposition 1a, namely, a Nash equilibrium of the game $\{N, \mathbf{A}, \boldsymbol{\pi}\}$ is also a consistent evidential equilibrium of the game $\{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$ with a suitable choice of social projection functions, $\mathbf{P}$.

Remark 5 *(Predictive content of evidential equilibria): As one would expect, the outcome in an evidential equilibrium depends critically on the social projection function (SPF). The four cases depicted above consider a range of social projection functions in the prisoners' dilemma game. The evidence, reviewed in Section 2 shows that when players consider others as like-minded and use evidential reasoning then there is good support for the identity SPF (Definition 6). In the prisoners' dilemma game, this corresponds to Case 1 above. However, there could be a mixture of players, some who use evidential reasoning while others use causal reasoning. Furthermore, beliefs of some players could turn out to be fulfilled, perhaps because they are particularly canny or lucky, while in other cases, beliefs may not be fulfilled. Using subsection 2.2, the weight of the evidence indicates a cooperation rate of at least 50% in the prisoners' dilemma game. When players are randomly matched and about half cooperate while the other half defect, the beliefs of some players are fulfilled but those of others are not. Given the observed heterogeneity of players, we consider a range of SPFs. However, the predictive content of evidential equilibrium for simple games with low dimensional choices is likely to arise from a mixture of players who use causal and evidential reasoning and the latter use the identity social projection function.*

5 Oligopoly games

We reconsider several classical models from industrial organization in the light of evidential reasoning; in particular the monopoly, competitive, Cournot, Bertrand and Stackelberg models. For ease of reference, we first give the classical formulation of these models under causal reasoning. We then reconsider them from the perspective of evidential reasoning and finally give the empirical evidence on such games.

We consider a market for a single homogeneous good. We shall assume a competitive market on the consumers' side, i.e., no consumer has any market power and they do not collude. Classically, this enables us to assume that each consumer is a price taker and, hence, that the unit price is given by an inverse-demand function, $P(Q)$; which we further assume to be a strictly decreasing function of the total industrial output of that good, Q. The total industrial output, Q, is produced by a fixed number of firms, n. Let q_i be the output of firm i. Then

$$Q = \sum_{i=1}^{n} q_i. \tag{4}$$

For simplicity we shall take $P(Q)$ to be linear,

$$P(Q) = A - aQ,\ a > 0,\ A > 0, \tag{5}$$

and we shall take the unit production cost of firm i to be a constant, c_i, $i = 1, 2, ..., n$, where

$$0 \leq c_1 \leq c_2 \leq ... \leq c_n < A. \tag{6}$$

Thus, the variable cost of firm i is $c_i q_i$. Assuming zero fixed costs, the total cost of firm i is $c_i q_i$. Hence, the profit of firm i is

$$\pi_i = (P - c_i)\, q_i,\ i = 1, 2, ..., n. \tag{7}$$

In the light of (4) and (5), the profit of firm i, (7), can be written in the useful form

$$\pi_i\left(q_i, \mathbf{q}_{-i}\right) = \left(A - c_i - a \sum_{j \neq i} q_j\right) q_i - a q_i^2,\ i = 1, 2, ..., n. \tag{8}$$

5.1 Classical oligopoly models under causal reasoning

In this subsection, we consider the textbook case where all players use causal reasoning.

5.1.1 Perfect competition

Under perfect competition, the market price, P^*, equals the minimum marginal cost, c_1:

$$P^* = c_1. \tag{9}$$

Hence, from (5), total output is

$$Q^* = \frac{A - c_1}{a}. \tag{10}$$

Note that outcomes (9)-(10) are not consistent with a Nash equilibrium when outputs (q_i) are the decision variables. However, they are consistent with a Nash equilibrium in the Bertrand game where prices, p_i, are the decision variables.

5.1.2 Monopoly

Suppose that a benevolent planner gives all rights of production and use of technology to a single profit maximizing monopolist. The chosen monopolist would use the lowest cost technology, resulting in the profit function

$$\Pi(Q) = (A - c_1)\, Q - a Q^2. \tag{11}$$

Maximizing (11) with respect to Q gives monopoly output[22]

$$Q^* = \frac{A - c_1}{2a}. \tag{12}$$

---

[22] By completing the square we get $\Pi(Q) = \frac{(A-c_1)^2}{4a} - a\left(Q - \frac{A-c_1}{2a}\right)^2$ which, clearly, has the unique maximum $Q = \frac{A-c_1}{2a}$.

5.1.3 Cournot oligopoly

Given the vector of outputs of all other firms, $\mathbf{q}_{-i}$, it easily follows from (8) that the profit maximizing output for firm i is

$$\widetilde{q}_i\left(\mathbf{q}_{-i}\right)=\frac{A-c_i}{2a}-\frac{1}{2}\sum_{j\neq i}q_j. \tag{13}$$

The Nash equilibrium, $\mathbf{q}^*$, also known as the Cournot equilibrium, is the fixed point of the function $\widetilde{\mathbf{q}}\left(\mathbf{q}\right)$. This can easily be found to be[23]

$$q_i^*=\frac{A+\sum_{j\neq i}c_j-nc_i}{\left(n+1\right)a}, i=1,2,...,n. \tag{14}$$

5.1.4 A Stackelberg leader-follower model

Unlike all the games considered above, which are single stage games, this is a two-stage game. There is a total of n firms. Firms $i=1,2,...,m$ are the followers while firms $i=m+1,m+2,...,n$ are the leaders. The followers choose their outputs, given the outputs of the leaders. When the leaders choose their outputs, they correctly anticipate the output choices of the followers.

To simplify the exposition, we concentrate on the case of equal unit costs:

$$0\leq c_1=c_2=...=c_n=c<A. \tag{15}$$

We rewrite the (reaction) function of follower i as

$$\widetilde{q}_i\left(\mathbf{q}_{-i}\right)=\frac{A-a\sum_{j=m+1}^{n}q_j-c}{2a}-\frac{1}{2}\sum_{j=1,j\neq i}^{m}q_j,\ i=1,2,...,m. \tag{16}$$

The Nash equilibrium for the followers, in the subgame determined by the leaders' outputs $q_{m+1},q_{m+2},...,q_n$, can be obtained by reinterpreting (14):

$$q_i^*=\frac{A-a\sum_{j=m+1}^{n}q_j-c}{\left(m+1\right)a}, i=1,2,...,m. \tag{17}$$

We could go on to calculate the subgame perfect equilibrium of this game. But of more interest to us is the case where the leaders maximize their joint profit, which they share

[23] The Cournot equilibrium, $\mathbf{q}^*$, must satisfy $q_i^*=\frac{A-c_i}{2a}-\frac{1}{2}\sum_{j\neq i}q_j^*$, $i=1,2,...,n$. Since $\sum_{j\neq i}q_j=Q-q_i$, this can be written as $q_i^*=\frac{A-c_i}{a}-Q^*$, $i=1,2,...,n$. Summing from 1 to n, and rearranging, gives $Q^*=\frac{nA-\sum_{j=1}^{n}c_j}{(n+1)a}$ and, hence, $q_i^*=\frac{A+\sum_{j\neq i}c_j-nc_i}{(n+1)a}$, $i=1,2,...,n$.

equally, correctly anticipating the reaction of the followers as given by (17). The resulting outputs are:

$$q_i^* = \frac{A-c}{2(m+1)a}, i = 1, 2, ..., m \text{ (Followers)}, \tag{18}$$

$$q_i^* = \frac{A-c}{2(n-m)a}, i = m+1, m+2, ..., n \text{ (Leaders)}. \tag{19}$$

From (17) we see that the followers, naturally, condition their outputs on that of the leaders. Hence, the leaders, when taking their decisions, anticipate the effect that their actions will have on the followers. This is entirely consistent with causal reasoning (recall footnote 13). Note that (18)-(19) form a Nash equilibrium, in fact a subgame perfect Nash equilibrium[24], only when $n = m + 1$ (*one* leader).

5.2 Evidential reasoning

In this subsection we focus on the consequences of evidential reasoning for the producers. But we shall assume that all consumers use causal reasoning, i.e., each consumer regards every other consumer and every firm as an outgroup member (recall Definition 5). We also assume that each firm regards each consumer as an outgroup member.[25] This also allows us to continue to assume that the market demand curve is given by (5).[26]

We now describe an evidential equilibrium, $\mathbf{q}^*$, with the following properties. Suppose that firm i is considering a deviation, q_i, from q_i^*. Firm i reasons as follows. "If I am tempted to deviate by an amount $q_i - q_i^*$ and if I believe that my rival, firm j, $j \neq i$, is like-minded, then the rival is probably also tempted to deviate by an amount $q_j - q_j^* = \lambda_{ij}(q_i - q_i^*)$"; the interpretation of λ_{ij} is given below. We formalize such reasoning by the following social projection function (Definition 2):

$$P_{ij}(q_j|q_i) = 1 \Leftrightarrow q_j - q_j^* = \lambda_{ij}(q_i - q_i^*), \; j \neq i. \tag{20}$$

The social projection specified in (20) is quite general. By specifying like-mindedness of other players in this manner, it nests several subcases, as we show below. At this point, the reader may wish to also think about Remark 5. The generality of (20) should not be

[24] I.e., Nash in the whole game and Nash in each subgame conditional on the output chosen by the leader.

[25] Thus if C is the set of consumers and F is the set of firms, then each is an outgroup relative to the other (Definition 2e).

[26] If we allowed consumers to use non-causal reasoning, then a single consumer could reason as follows "If I cut my demand, then probably each like-minded consumer would also cut his demand. The aggregate result would be a reduction in price for all of us". Consumers would then be able to collude. The consequence would be that we would no longer have an oligopoly model (as classically defined) but a bargaining model. While this is very interesting, it lies beyond the scope of this paper and, in fact, deserves a paper on its own.

taken to mean that the predictive content of the evidential reasoning model of oligopoly is empty. Rather, as in prisoner's dilemma games, individuals display a wide variation in choices when they are asked to play oligopoly game (see Section 5.3 below). Variations in the parameter λ_{ij} in (20) offer a parsimonious way of capturing this heterogeneity.

In particular, changes in λ_{ij} may be thought of as capturing how different degrees of like-mindedness. For instance, perfect like-mindedness, $\lambda_{ij} = 1$, gives rise to the identity social projection function (Definition 6). The other extreme arises when no like-mindedness is perceived by firms, as in models of causal reasoning. This corresponds to $\lambda_{ij} = 0$. Intermediate cases of like-mindedness correspond to values $0 < \lambda_{ij} < 1$ and to $\lambda_{ij} < 0$ and possibly reflect different degrees of uncertainty that players have about their opponents. The distribution of values of λ_{ij} in any population is eventually an empirical question that cannot be answered ex-ante.

Lemma 2 *: (a) Given the social projection functions (20), the unique evidential equilibrium (Definition 10),* $\mathbf{q}^*$*, is characterized by the following set of simultaneous linear algebraic equations*

$$\begin{bmatrix} 2+\sum_{j\neq 1}\lambda_{1j} & 1 & ... & 1 \\ 1 & 2+\sum_{j\neq 2}\lambda_{2j} & ... & 1 \\ ... & ... & ... & ... \\ 1 & 1 & ... & 2+\sum_{j\neq n}\lambda_{nj} \end{bmatrix} \begin{bmatrix} q_1^* \\ q_2^* \\ ... \\ q_n^* \end{bmatrix} = \begin{bmatrix} \frac{A-c_1}{a} \\ \frac{A-c_2}{a} \\ ... \\ \frac{A-c_n}{a} \end{bmatrix} \tag{21}$$

(b) Furthermore, $\mathbf{q}^*$ *is a mutually consistent vector of strategies (Definition 11) and, hence, a consistent evidential equilibrium.*
(c) Conversely, given any vector of outputs, $\mathbf{q}^*$*, satisfying* $q_i^* > 0$ *and* $\sum_{i=1}^{n} q_i^* \leq \frac{A-c_1}{a}$*, there exits a profile of social projection of the form (20) such that* $\mathbf{q}^*$ *is a consistent evidential equilibrium.*

Proof of Lemma 2: (a) Substituting q_j from (20) into (8) gives

$$\pi_i\left(q_i, \mathbf{P}_i\left(.|q_i\right)\right) = \left\{A - c_i - a\sum_{j\neq i}\left[q_j^* + \lambda_{ij}\left(q_i - q_i^*\right)\right]\right\} q_i - aq_i^2,\ i = 1, 2, ..., n, \tag{22}$$

which, after simplification, gives

$$\pi_i\left(q_i, \mathbf{P}_i\left(.|q_i\right)\right) = \left(A - c_i + aq_i^*\sum_{j\neq i}\lambda_{ij} - a\sum_{j\neq i}q_j^*\right) q_i - a\left(1 + \sum_{j\neq i}\lambda_{ij}\right) q_i^2,\ i = 1, 2, ..., n. \tag{23}$$

(23) shows how a player who uses the heuristic of evidential reasoning translates an essentially strategic problem into a decision theoretic problem. Maximizing (23) with respect

to q_i gives the optimal (pure) strategy for firm i (Definition 9), given his social projection function (20):

$$q_i = \frac{A - c_i + aq_i^* \sum_{j \neq i} \lambda_{ij} - a \sum_{j \neq i} q_j^*}{2a \left(1 + \sum_{j \neq i} \lambda_{ij}\right)},\ i = 1, 2, ..., n. \tag{24}$$

Setting $q_i = q_i^*$, $i = 1, 2, ..., n$ and simplifying gives the following set of simultaneous linear algebraic equations,

$$\left(2 + \sum_{j \neq i} \lambda_{ij}\right) q_i^* + \sum_{j \neq i} q_j^* = \frac{A - c_i}{a},\ i = 1, 2, ..., n, \tag{25}$$

which can be written in the matrix form

$$\begin{bmatrix} 2 + \sum_{j \neq 1} \lambda_{1j} & 1 & ... & 1 \\ 1 & 2 + \sum_{j \neq 2} \lambda_{2j} & ... & 1 \\ ... & ... & ... & ... \\ 1 & 1 & ... & 2 + \sum_{j \neq n} \lambda_{nj} \end{bmatrix} \begin{bmatrix} q_1^* \\ q_2^* \\ ... \\ q_n^* \end{bmatrix} = \begin{bmatrix} \frac{A - c_1}{a} \\ \frac{A - c_2}{a} \\ ... \\ \frac{A - c_n}{a} \end{bmatrix} \tag{26}$$

(b) From (20) we see that $P_{ij}\left(q_j | q_i^*\right) = 1 \Leftrightarrow q_j = q_j^*$. In effect, when firm i produces the output q_i^* it assigns perfect certainty of beliefs that firm j will produce q_j^*. Ex-post, firm i finds that firm j indeed did produce an output level q_j^*, thus, vindicating its ex-ante beliefs. Hence, $\mathbf{q}^*$ is a mutually consistent vector of strategies and, hence, a consistent evidential equilibrium.

(c) Rewrite (25) in the form

$$\sum_{j \neq i} \lambda_{ij} = \frac{A - c_i}{aq_i^*} - 2 - \frac{1}{q_i^*} \sum_{j \neq i} q_j^*,\ i = 1, 2, ..., n. \tag{27}$$

(27) has many solutions, for example

$$\begin{aligned} \lambda_{ij} &= \lambda_i,\ i, j = 1, 2, ..., n,\ j \neq i, \text{ where} \\ \lambda_i &= \frac{A - c_i}{(n-1)\, aq_i^*} - \frac{2}{n-1} - \frac{1}{(n-1)\, q_i^*} \sum_{j \neq i} q_j^*,\ i = 1, 2, ..., n.\ \blacksquare. \end{aligned}$$

In subsection 5.1 we showed how under causal reasoning one may obtain various market outcomes such as perfect competition, monopoly, Cournot oligopoly, and Stackelberg leader-follower. As an application of Lemma 2 we now show how one may obtain the same market outcomes under evidential reasoning by choosing suitable values for λ_{ij}, $j \neq i$, in (21). This is followed in section 5.3 by a consideration of the empirical evidence.

5.2.1 Monopoly

Setting $c_1 = c_2 = ... = c_n$ and $\lambda_{ij} = 1$, for all i, j and $i \neq j$, in (21) gives $q_i^* = \frac{A - c_1}{2na}$, $i = 1, 2, ..., n$. Hence, the total output level $\sum_{i=1}^{n} q_i^* = Q^* = \frac{A - c_1}{2a}$, which is identical to the monopoly output level given in (12). In this case, the social projection functions for the producers are identity social projection functions on the set of all producers (Definition 6). The set of firms form a perfect ingroup (Definition 7). They behave harmoniously (or cooperatively) towards each other (see subsection 3.3). Thus, if all firms believe that others are like-minded and they use the identity social projection function then the aggregate output maximizes joint profits of the group.

5.2.2 Cournot oligopoly

Setting $\lambda_{ij} = 0$, for all i, j and $i \neq j$, in (21) gives $q_i^* = \frac{A + \sum_{j \neq i} c_j - nc_i}{(n+1)a}$, $i = 1, 2, ..., n$, which is identical to the Cournot output levels given in (14). Here each firm regards the other firm as an outgroup member (Definition 5). Thus, each firm regards every other player (whether consumer or producer) as an outgroup member. Hence, every firm uses causal reasoning (Definition 3).

5.2.3 Perfect competition

Setting $c_1 = c_2 = ... = c_n$ and $\lambda_{ij} = -\frac{1}{n-1}$, for all i, j and $i \neq j$, in (21) gives $\sum_{i=1}^{n} q_i^* = Q^* = \frac{A - c_1}{a}$, in agreement with (10). Here each firm regards every other firm as an ingroup member and the set of firms forms an ingroup (Definition 5). We may call this a *competitive ingroup* and the resulting social projection functions *competitive social projection functions.* This is in line with the ideas considered in subsection 3.3 and, in particular, is an illustration of the contrast effect.

5.2.4 Stackelberg leader-follower model

Consider the case $c_1 = c_2 = ... = c_n$ and let

$$\begin{aligned} \lambda_{ij} &= 0,\, i = 1, 2, ..., m,\, j = 1, 2, ..., n,\, i \neq j, \\ \lambda_{ij} &= -\frac{n-m}{m+1},\, i = m+1, m+2, ..., n,\, j = 1, 2, ..., m, \\ \lambda_{ij} &= 1,\, i = m+1, m+2, ..., n,\, j = m+1, m+2, ..., n,\, i \neq j. \end{aligned}$$

Here the social projection function of each of the leaders (firms $m+1, m+2, ..., n$) is an identity social projection function on the set of leaders (Definition 6). Hence, the leaders form a perfect ingroup (Definition 7). The social projection functions for the followers (firms $1, 2, ..., m$) are all constant, hence all the followers use causal reasoning (Definition 3). Each follower regards each leader as an outgroup member. The leaders form an outgroup relative to the followers and each leader regards each follower as an ingroup member (Definition 5) although the degree of like mindedness they assign to followers ($\lambda_{ij} = -\frac{n-m}{m+1}$) is lower relative to that they assign to other leaders ($\lambda_{ij} = 1$). We may say that the leaders behave *collusively* towards each other but *competitively* towards the followers (the contrast effect, recall subsection 3.3).

Substitute the assumed values of λ_{ij} in (21) to get

$$\begin{aligned} q_i^* &= \frac{A-c}{2(m+1)a}, i=1,2,...,m \text{ (Followers)}, \\ q_i^* &= \frac{A-c}{2(n-m)a}, i=m+1,m+2,...,n \text{ (Leaders)}. \end{aligned}$$

Hence, the output levels of leaders and followers are identical to those under the Stackelberg case when players use causal reasoning as in (18) and (19). However, unlike the Stackelberg game of subsection 5.1.4, this version is a single-stage game.[27] The empirical evidence that we present below does show that in static duopoly games, players often choose outputs similar to the Stackelberg output level.

The following is a simple corollary of Lemma 2, which is of interest in its own right.

Corollary 3 *Suppose all firms are identical, so that $c_1 = c_2 = ... = c_n = c$ (say) and $\lambda_{ij} = \lambda$, $i, j = 1, 2, ..., n$, $j \neq i$. Then under the social projection functions (20):*
(a) The consistent evidential equilibrium, $\mathbf{q}^*$, is given by

$$q_i^* = \frac{A-c}{[n+1+(n-1)\lambda]a}, \ i=1,2,...,n.$$

(b) The profit of firm i is given by

$$\pi_i^* = \frac{[1+(n-1)\lambda](A-c)^2}{a[n+1+(n-1)\lambda]^2}, \ i=1,2,...,n.$$

(c) π_i^* is strictly increasing in λ in the range $-\frac{n+1}{n-1} < \lambda < 1$.
(d) π_i^* is maximized when $\lambda = 1$.
(e) In particular, as λ increases from $-\frac{1}{n-1}$ to 1, the profit (output) level of each firm increases (decreases) from the perfectly competitive, through the Cournot ($\lambda = 0$), to the fully collusive.

5.3 Empirical evidence from oligopoly games

In our theoretical model we have used a flexible social projection function in (20). We showed how alternative values of λ in (20) were able to produce, in a static game, the output levels under alternative market forms such as monopoly, Cournot duopoly, Stackelberg leader-follower game, perfect competition etc. The aim of this empirical section is to show

[27] For the special case $n = 2$, $m = 1$ (one leader and one follower) we get: $q_1^* = \frac{A-c}{4a}$ (follower), $q_2^* = \frac{A-c}{2a}$ (leader), $Q^* = \frac{3(A-c)}{4a}$ (total output).

that one indeed does observe a wide range of market outcomes when a duopoly game is played with experimental subjects.

In the early experiments on Cournot markets by Fouraker and Siegel (1963), and the experiments that followed for several decades, it was usual to present a *profit table* (PT). The PT was typically based on linear demand and linear cost curves in symmetric, homogenous goods duopolists. The PT listed the outputs of each firm on the two margins while individual cells of the table contained the corresponding profits of both firms. The PT was often supplemented by a *profit calculator* (PC) which allowed each experimental subject in their role as a firm to calculate the profit for a given pair of quantities chosen by both firms. In recent years, several experiments also give to the subjects a *best response option* (BRO) which tells them their profit maximizing quantity for any quantity chosen by the other player.

The extra information provided (PT, PC, BRO) arguably alters the nature of the problem by suggesting a particular frame and solution. Requate and Waichman (2011) find that there is substantially more collusion (corresponding to $\lambda = 1$ in (20)) in PT and PC treatments as compared to BRO. They find, in a static duopoly experiment, that the collusive outcome is reached at least once in the 20 rounds, in 62%, 78% and 29% of the markets, respectively, in the PT, PC and BRO treatments. The theoretical outcome of the Cournot-Nash equilibrium is, therefore, not confirmed in many cases.

Several papers claim to find support for the Cournot-Nash equilibrium under random matching of opponents while finding that there is greater collusion under fixed matching of players.[28] Consider a representative paper by Huck et al. (1999), which uses symmetric firms and linear demand curves. For the demand and cost curves used, the Cournot-Nash outcome is for each firm to produce an output of 8. The authors report data for round 9 of play under random matching, as the most supportive of their hypothesis (see Table 5 in their paper); these results are given in Table 7.[29]

Output level	6	7	8	9	greater than 10
% of subjects choosing	12	21.5	35.5	14.5	14

Table 7: Distribution of output levels in Huck et al. (1999)

The mean quantity is close to the Cournot-Nash output level of 8. However, there is substantial variation in the output levels and about 65% of individuals do not choose the Cournot output level.

[28] The interested reader can consult the bibliography in Requate and Waichman (2011).

[29] In these experiments, outputs of both firms vary between 3-15. The Cournot-Nash outcome is for each firm to produce an output of 8. Since both firms are known to the experimental subjects to be symmetric, they must know that the solution lies on the diagonal of a relatively small matrix. Profits of each firm in the PT drop off sharply for output levels equal to or higher than 10 or less than an output of 3. These arguably leaves only 7 levels of output to choose from: 4,5,6,7,8,9.

Rassenti et al. (2000) use an asymmetric Cournot game in which firms have different marginal costs. Importantly, firms are not told the marginal costs of their opponents or any probability distribution over them. In this sense, there is true uncertainty, an area where evidential reasoning would seem to have the most bite. The game is played over 75 rounds to allow for substantial learning possibilities. The main finding is that while total output is above, but close to, the Cournot-Nash solution, the individual levels of output chosen by the firms are quite different from the Cournot-Nash solution. The results, in this sense, are similar to those in Huck et al. (1999), however, the authors take this as a refutation rather than a confirmation of the Cournot-Nash equilibrium.

Bosch-Domènech and Vriend (2003) report results from the last two rounds of a 22 round duopoly experiment in which, in each round, two firms simultaneously choose outputs. They find that the output levels are widely distributed over a range that includes the monopoly output level and the perfectly competitive level. Table 8 summarizes information that is extracted from their paper. The three treatments, *easy*, *hard* and *hardest*, differ in terms of the time within which firms had to choose their outputs and the level of information provided.[30] The cells in Table 8 report the approximate percentage of output chosen by the experimental subjects under each column head. It is clear from the evidence that the Cournot output level is not particularly salient relative to the others. Further, the wide distribution of output levels even in rounds 21 and 22 of the experiment suggest that a flexible social projection function as in (20) is consistent with the evidence.

	Monopoly	Cournot	Walrasian	Others
Easy	38.89	33.33	0	27.78
Hard	11.11	16.67	11.11	61.11
Hardest	8.33	14.28	16.67	60.72

Table 8: Percentage of outputs corresponding to various market levels

When players choose their quantities simultaneously, can they observe each other's body language, talk to each other or send other kinds of messages? The Cournot-Nash equilibrium is agnostic about pre-play communication; such features are simply not a part of the game. Waichman et al. (2010) find that pre-play communication increases the degree of collusion in the Cournot game. Between 91% and 100% of the markets achieve collusion in at least one round of the experiment when pre-play communication is allowed. This also seems consistent with evidential reasoning. Pre-play communication may increase the players' beliefs that they are dealing with like-minded players, hence, facilitating the use

[30]For instance, in the easy treatment, a PT is provided but not in the other treatments. In the hardest treatment, firms are not even told of the exact functional form of the linear demand curve (only that it is downward sloping) whereas in the easy treatment firms know the exact demand curve.

of evidential reasoning.[31]

Duersch et al. (2010) document systematic departures from a Cournot-Nash equilibrium. They consider a linear-demand, linear-cost Cournot game with the Cournot-Nash quantity, $q_i^* = 36$. Computers play one of several well known strategies including best response against human subjects who are not aware of the computers' strategy over 40 rounds. Again, by creating uncertainty about what others will do, this situation is quite relevant to the domain of evidential reasoning.

Mean quantities chosen by computers (34.39) are always lower than mean quantities chosen by humans (47.95). Human subjects choose quantities that are much greater than the Cournot-Nash levels and, in some cases, approach the Stackelberg leader output of 54. In particular, when computers are programmed to play a best response with some small noise, in three different treatments, subjects choose the output levels 51.99, 48.67, and 49.18, while computers choose 32.05, 35.02, 31.67. Thus, human subjects show systematic (upward) departures from the Cournot-Nash level, even approaching the Stackelberg levels.

5.4 Summary of the evidence in oligopoly games

Among the classical oligopoly models (subsection 5.1, where all players use causal reasoning), $\mathbf{q}^*$ is a Nash equilibrium only for the case of the Cournot model (subsection 5.1.3).[32] However, under evidential reasoning (subsection 5.2), $\mathbf{q}^*$ is always a consistent evidential equilibrium.

Turning now to the empirical evidence (subsection 5.3), the claim that under random matching, experimental subjects' behavior robustly conforms to a Cournot-Nash equilibrium is not supported by the evidence. One observes a wide and rich range of behaviors that are often collusive and range all the way up to the choice of quantities in the Stackelberg case. The results are consistent with a range of social projection functions used (corresponding to alternative values of λ in (20)) by experimental subjects who use evidential reasoning. The collusive outcome is played by a significant percentage of experimental subjects, which is consistent with the identity social projection function.

[31] Recall from the discussion of the work by Darai and Grätz (2010) above that in the prisoner's dilemma game with pre-play communication and handshakes, the cooperation levels are also similarly enhanced.

[32] Barring some very special cases, as indicated in subsection 5.1.

6 Evidential equilibrium in static games of incomplete information

6.1 Motivation

A player might be unsure about the payoff functions or the action sets of the other players. In addition to these typical concerns, a player may be unsure as to whether some of the other players are ingroup or outgroup members, as illustrated in Example 5, below. Our formulation, in subsection 6.3, below, allows for a consideration of both these concerns.

Example 5 *(Prisoners' dilemma game of incomplete information): Suppose that you, player 1, are particularly well disposed to cooperate with others and you spend your spare time working gratis for various charities. You have been asked to play a prisoners' dilemma game with an opponent, player 2. You believe that there is a probability* $p \geq 0$ *that player 2 was fired from his previous job on the grounds of being an atrocious team player who also never privately contributes to charity (type 1 of player 2) and a probability* $1 - p$ *that he was never fired from his job and contributes to charity (type 2 of player 2). Should player 1 cooperate?*
This is an example of a static game of incomplete information. Player 2 has more than one possible type that is privately known to player 2 but not to player 1. Player 1 might reason as follows: "There is a probability p *that player 2 is not like-minded, so in this case, I may consider him as an outgroup member. In particular, I expect such type of a player to always choose the strategy defect. But there is a probability* $1 - p$ *that player 2 is completely like-minded so I may take my own propensity to cooperate or defect as having perfect diagnostic significance in this case".*[33] *Notice that in this case, player 1 uses causal reasoning towards type 1 of player 2 and evidential reasoning towards type 2 of player 2 and, in particular, assigns the identity social projection function towards the type 2 player.*

The above example indicates that the concept of evidential reasoning can also be used for static games of incomplete information. We give, below, a formal development of evidential reasoning in the context of static games of incomplete information.

6.2 Preliminary considerations

In standard accounts of game theory, the terms "static games of incomplete information" and "static Bayesian games" are used as synonymous. However, we wish to distinguish between them. In particular, in our formulation of static games of incomplete information,

[33] Or one may have a richer setting in which player 1 believes that the type 2 of player 2 is like-minded with probability $q \neq 1 - p$ and not like-minded with probability $1 - q$. Our setup allows for this feature although, for simplicity, we shall assume the case stated in the text.

we do not commit to common priors, Bayesian updating or the expected utility form of payoffs. Adding these changes a static game of incomplete information into a static Bayesian game as commonly understood in standard game theory.

Our motivation for introducing evidential reasoning is entirely empirical. Given this attitude, it does not appear appropriate to introduce evidential reasoning into a theory where three of the fundamental assumptions, namely, common priors, Bayesian updating and expected utility are not empirically supported by the evidence.[34] Of course, a fully developed theory should include substitutes for these concepts.[35] However, this lies beyond the scope of this paper.

6.2.1 Static games of incomplete information

As in section 3, $N = \{1, 2, ..., n\}$ is the set of players. However, now each player $i \in N$ can be one of several types, given by $T_i = \{t_{i1}, t_{i2}, ..., t_{in_i}\}$, with typical element denoted by t_i. The Cartesian product of types, or type space, is given by the set $\mathbf{T} = \times_{i=1}^{n} T_i$, and a typical member is denoted by $\mathbf{t} = (t_1, t_2, ..., t_n) \in \mathbf{T}$. For any player, i, the set of types of all other players is $\mathbf{T}_{-i} = \times_{j=1, j\neq i}^{n} T_i$, and a typical member is denoted by $\mathbf{t}_{-i} = (t_1, t_2, ..., t_{i-1}, t_{i+1}, ..., t_n)$. Each player has incomplete information about the exact type of the opponents. Following Harsanyi's well known device, we assume that nature moves first and draws some vector of types $\mathbf{t} = (t_1, t_2, ..., t_n) \in \mathbf{T}$. To each player $i \in N$, nature privately reveals the corresponding type $t_i \in T_i$ that was drawn but does not give this information to any other player. This changes a game of incomplete information into a game of imperfect information.

The probability assigned by player i that the type vector of other players is $\mathbf{t}_{-i}$, conditional on his own type being t_i, is given by $p_i (\mathbf{t}_{-i} \mid t_i)$. When types are independent then $p_i (\mathbf{t}_{-i} \mid t_i)$ does not depend on t_i and $p_i (\mathbf{t}_{-i} \mid t_i) = p_i (\mathbf{t}_{-i})$. Let $\mathbf{p} = \{p_i (\mathbf{t}_{-i} \mid t_i) : i \in N, \mathbf{t} \in \mathbf{T}\}$ be the family of these beliefs.

The set of actions of a player $i \in N$, of type $t_i \in T_i$, is given by $A_i (t_i)$. We assume $A_i (t_i)$ to be finite. Let $\mathbf{A} (\mathbf{t}) = \times_{i=1}^{n} A_i (t_i)$. Denote by $\mathbf{A} = \{A_i (t_i) : i \in N, \mathbf{t} \in \mathbf{T}\}$, the family of all action sets of all types of players.

The set of probability distributions over $A_i (t_i)$ is given by $\Delta_i (t_i)$. For any player i, a strategy of type $t_i \in T_i$ is given by $\sigma_i (., t_i) \in \Delta_i (t_i)$, thus, strategies are type-contingent. $\sigma_i (a_i, t_i)$ is the probability that type $t_i \in T_i$, of player i, takes the action $a_i \in A_i (t_i)$. Let $\mathbf{\Delta} = \{\Delta_i (t_i) : i \in N, \mathbf{t} \in \mathbf{T}\}$ be the family of all the strategies of all the types of players.

[34] See, for instance, Kahneman and Tversky (2000) and Camerer (2003).

[35] For instance, prospect theory or rank dependent theory in place of expected utility; judgement heuristics in place of Bayes' rule etc.

Definition 13 *(Payoff functions under incomplete information): Let*

$$\boldsymbol{\sigma}_i = \left(\sigma_i\left(.,t_{i1}\right), \sigma_i\left(.,t_{i2}\right), ..., \sigma_i\left(.,t_{in_i}\right)\right)$$

be a vector of strategies for player i, one for each type of player i. Let $\boldsymbol{\sigma} = \left(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, ..., \boldsymbol{\sigma}_n\right)$ be a profile of such vectors, one for each player and let $\boldsymbol{\sigma}_{-i} = \left(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, ..., \boldsymbol{\sigma}_{i-1}, \boldsymbol{\sigma}_{i+1}, ..., \boldsymbol{\sigma}_n\right)$ be the profile of strategy vectors of the players other than player i. The payoff to player i of type t_i as a result of that player adopting the strategy $\sigma_i\left(.,t_i\right)$ and the other players adopting the profile $\boldsymbol{\sigma}_{-i}$ is written as

$$U_i\left(\sigma_i\left(.,t_i\right), \boldsymbol{\sigma}_{-i}\right). \tag{28}$$

(28) captures the idea that player i knows his own type, t_i, but not, necessarily, that of the other players. Hence, his payoff will depend on his own strategy, $\sigma_i\left(.,t_i\right)$, and, in general, will also depend on the strategy adopted by each type of each of the other players, as captured by $\boldsymbol{\sigma}_{-i}$. In addition, the payoff to player i, U_i, will depend on the type space, $\mathbf{T}$, and also on the assessment by player i of the probability, $\left\{p_i\left(\mathbf{t}_{-i} \mid t_i\right) : \mathbf{t}_{-i} \in \mathbf{T}_{-i}\right\}$, of each type of each of the other players, given knowledge of his own type t_i. Let $\mathbf{U} = \left\{U_i : i \in N\right\}$ be the family of the payoff functions of the players.

Definition 14 *: A static game of incomplete information is given by $\mathbf{G} = \langle N, \mathbf{T}, \mathbf{p}, \mathbf{A}, \boldsymbol{\Delta}, \mathbf{U}\rangle$, where N, $\mathbf{T}$, $\mathbf{p}$, $\mathbf{A}$, $\boldsymbol{\Delta}$, $\mathbf{U}$ have all been introduced above.*

Definition 15 *: $\sigma_i^*\left(.,t_i\right)$ is a best reply for player i of type t_i to the profile of strategy vectors $\boldsymbol{\sigma}_{-i}$ adopted by all the types of the other players, if $U_i\left(\sigma_i^*, \boldsymbol{\sigma}_{-i}\right) \geq U_i\left(\sigma_i, \boldsymbol{\sigma}_{-i}\right)$ for each $\sigma_i \in \Delta_i$ (recall Definition 13).*

Definition 16 *: A profile of strategy vectors, $\boldsymbol{\sigma}^*$, is a Nash equilibrium under incomplete information if, for each i, $i = 1, 2, ..., n$, and each $t_i \in T_i$, $\sigma_i^*\left(.,t_i\right)$ is a best reply to $\boldsymbol{\sigma}_{-i}^*$ (recall Definitions 13 and 15).*

Note that in computing a Nash equilibrium of the game of incomplete information, $\mathbf{G} = \langle N, \mathbf{T}, \mathbf{p}, \mathbf{A}, \boldsymbol{\Delta}, \mathbf{U}\rangle$, one needs to specify an optimal strategy for each type of each player.

6.2.2 Static Bayesian games

Definition 17 *: A static Bayesian game is a static game of incomplete information (Definition 14), $\mathbf{G} = \langle N, \mathbf{T}, \mathbf{p}, \mathbf{A}, \boldsymbol{\Delta}, \mathbf{U}\rangle$, where the family of beliefs, $\mathbf{p} = \left\{p_i\left(\mathbf{t}_{-i} \mid t_i\right) : i \in N, \mathbf{t} \in \mathbf{T}\right\}$,* *of the players about the types of the other players is derived, using Bayes law, from a common prior, $p(\mathbf{t})$, which is common knowledge to all players:*

$$p_i(\mathbf{t}_{-i} \mid t_i) = \frac{p(\mathbf{t})}{\sum_{\mathbf{t}_{-i} \in \mathbf{T}_{-i}} p(\mathbf{t}_{-i}, t_i)}, \tag{29}$$

and the payoff functions (28) take the expected utility form (assuming uncorrelated strategies):

$$U_i(\sigma_i, \boldsymbol{\sigma}_{-i}) = \sum_{\mathbf{t}_{-i} \in \mathbf{T}_{-i}} \sum_{\mathbf{a} \in \mathbf{A}(\mathbf{t})} p_i(\mathbf{t}_{-i} \mid t_i) \Pi_{j=1}^{n} \sigma_j(a_j, t_j) u_i(\mathbf{a}, t_i, \mathbf{t}_{-i}). \tag{30}$$

Definition 18 *(Bayesian Nash equilibrium): A Nash equilibrium of a static Bayesian game (Definition 17) is called a Bayesian Nash equilibrium.*

Note that for a Bayesian Nash equilibrium, the strategy of each type of each player must be a best reply to the profile of strategy vectors of other player, just like a Nash equilibrium of a static game of incomplete information. But, in addition, the beliefs of each type of each player must be derived from a common prior using Bayes' law (29). And, furthermore, the payoff of each type of each player takes the expected utility form (30).

6.3 Static evidential games of incomplete information

We now consider the implications of evidential reasoning for static games of incomplete information. We begin by defining *social projection functions* (SPF) for such games. The main difference is that the social projection function, SPF (recall Definition 2) is now type-dependent. In particular, type t_i of player i must now form beliefs about the action to be taken by *each type of each of other players*, as in the example of subsection 6.1.

Definition 19 *: A social projection function for type t_i of player i onto type t_j of player j, $i \neq j$ (SPF for short) in a static game of incomplete information is a mapping $\mathbf{P}_{it_i,jt_j} : \Delta_i(t_i) \to \Delta_j(t_j)$, that assigns to each strategy, $\sigma_i(.,t_i) \in \Delta_i(t_i)$, for type t_i of player i, a strategy, $\sigma_j^e(.,t_j) \in \Delta_j(t_j)$, for type t_j of player j, $j \neq i$. If $\mathbf{P}_{it_i,jt_j}(\sigma_i(.,t_i)) = \sigma_j^e(.,t_j)$, $i \neq j$, then $\sigma_j^e(a,t_j)$ is the probability with which player i of type t_i believes player j of type t_j will play action $a \in A_j(t_j)$; which we write as $\mathbf{P}_{it_i,jt_j}(a|\sigma_i(.,t_i))$.*

Definition 20 *(Family of social projection functions): We call*

$$\mathbf{P} = \left\{\mathbf{P}_{it_i,jt_j} : t_i \in T_i, t_j \in T_j, i, j \in N, i \neq j\right\},$$

where $\mathbf{P}_{it_i,jt_j}$ are as in Definition 19, a family of social projection functions.

Definition 21 *(Anticipated strategies and anticipated payoffs): Let* $\mathbf{P}_{it_i,jt_j} : \Delta_i(t_i) \to \Delta_j(t_j)$ *be a social projection function for type* t_i *of player* i *onto type* t_j *of player* j*,* $j \neq i$ *(Definition 19). Let* $\sigma_i(.,t_i) \in \Delta_i(t_i)$ *be a strategy for type* t_i *of player* i*. Suppose* $\mathbf{P}_{it_i,jt_j}(\sigma_i(.,t_i)) = \sigma_j^e(.,t_j)$*. We then call* $\sigma_j^e(.,t_j)$ *the strategy of type* t_j *of player* j *anticipated by player* i*, when type* t_i *of player* i *uses his strategy* $\sigma_i(.,t_i)$*,* $i \neq j$*. We write* $\boldsymbol{\sigma}_j^e(\sigma_i(.,t_i)) = \left(\sigma_j^e(.,t_{j1}), \sigma_j^e(.,t_{j2}), ..., \sigma_j^e\left(.,t_{jn_j}\right)\right)$*,* $j \neq i$*, for the vector of strategies of each of the types of player* j *as anticipated by type* t_i *of player* i *when he uses his strategy* $\sigma_i(.,t_i)$*. We write*

$$\boldsymbol{\sigma}_{-i}^e(\sigma_i(.,t_i)) = \left(\boldsymbol{\sigma}_1^e(.), \boldsymbol{\sigma}_2^e(.), ..., \boldsymbol{\sigma}_{i-1}^e(.), \boldsymbol{\sigma}_{i+1}^e(.), ..., \boldsymbol{\sigma}_n^e(.)\right)$$

for the profile of strategy vectors that type t_i *of player* i*, when using his strategy* $\sigma_i(.,t_i)$*, anticipates all the types of all the other players will use. We write* $U_i\left(\sigma_i(.,t_i), \boldsymbol{\sigma}_{-i}^e(\sigma_i(.,t_i))\right)$ *for the payoff anticipated by player* i *of type* t_i*.*

As in static games of complete information, no player is required to be aware of the social projection functions of the other player, hence, players cannot strategically exploit the social projection functions of others.

Definition 22 *(Causal Reasoning): In a static game of incomplete information, we say that type* t_i *of player* i *uses causal reasoning if* $\mathbf{P}_{it_i,jt_j}(\sigma_i(.,t_i))$ *is independent of* $\sigma_i(.,t_i)$*, i.e., if* $\mathbf{P}_{it_i,jt_j}(\sigma_i(.,t_i)) = \mathbf{P}_{it_i,jt_j}(\sigma_i'(.,t_i))$ *for all* $\sigma_i(.,t_i), \sigma_i'(.,t_i) \in \Delta_i(t_i)$*, all* $t_j \in T_j$ *and all* $j \neq i$*.*

Definition 23 *(Evidential Reasoning): In a static game of incomplete information, we say that type* t_i *of player* i *uses evidential reasoning if it is not necessarily the case that* $\mathbf{P}_{it_i,jt_j}(\sigma_i(.,t_i)) = \mathbf{P}_{it_i,jt_j}(\sigma_i'(.,t_i))$ *for all* $\sigma_i(.,t_i), \sigma_i'(.,t_i) \in \Delta_i(t_i)$*, all* $t_j \in T_j$ *and all* $j \neq i$*.*

Definition 24 *: A static evidential game of incomplete information,* $\boldsymbol{\Gamma} = \langle \mathbf{G}, \mathbf{P} \rangle$*, is a static game of incomplete information,* $\mathbf{G}$ *(Definition 14), that is augmented with a family of social projection functions,* $\mathbf{P}$ *(Definition 20).*

Definition 25 *(Ingroups and Outgroups): Suppose that players use evidential reasoning.*
(a) Player i *of type* t_i *regards player* j $(j \neq i)$ *of type* t_j *as an outgroup member if* $\mathbf{P}_{it_i,jt_j}(\sigma_i(.,t_i))$ *is independent of* $\sigma_i(.,t_i)$*, i.e., if* $\mathbf{P}_{it_i,jt_j}(\sigma_i(.,t_i)) = \mathbf{P}_{it_i,jt_j}(\sigma_i'(.,t_i))$ *for all* $\sigma_i(.,t_i), \sigma_i'(.,t_i) \in \Delta_i(t_i)$*. Otherwise, type* t_i *of player* i *regards type* t_j *player* j $(j \neq i)$ *as an ingroup member.*
(b) Let $M \subset N$ *be a non-empty set of players. If every type of every player in* M *regards every type of every other player in* M *as an ingroup member, then* M *is an ingroup.*

(c) Let $L \subset N$ and $M \subset N$ be disjoint non-empty sets of players. Suppose every type of every player in L regards every type of every player in M as an outgroup member. Then we say that M is an outgroup relative to L.

Definition 26 *(Optimal strategies): An optimal strategy for type t_i of player i, $\sigma_i^*(.,t_i) \in \Delta_i(t_i)$, in the static evidential game of incomplete information, $\mathbf{\Gamma} = \langle N, \mathbf{T}, \mathbf{p}, \mathbf{A}, \mathbf{\Delta}, \mathbf{U}, \mathbf{P}\rangle$ (Definition 24), is one that maximizes the anticipated payoff, $U_i\left(\sigma_i(.,t_i), \boldsymbol{\sigma}_{-i}^e(\sigma_i(.,t_i))\right)$, for type t_i of player i, i.e., $U_i\left(\sigma_i^*(.,t_i), \boldsymbol{\sigma}_{-i}^e(\sigma_i^*(.,t_i))\right) \geq U_i\left(\sigma_i(.,t_i), \boldsymbol{\sigma}_{-i}^e(\sigma_i(.,t_i))\right)$, for all $\sigma_i(.,t_i) \in \Delta_i(t_i)$ (recall Definition 21).*

As in static games of complete information, each type t_i of each player i solves a decision theoretic problem conditional on the privately held beliefs, $\left\{\mathbf{P}_{it_i,jt_j} : j \in N - \{i\}, t_j \in T_j\right\}$. If the payoff functions are continuous and the choice space is compact then an optimal value for the anticipated payoff is assured.

Definition 27 *(Evidential equilibria and consistent evidential equilibria under incomplete information): Consider the static evidential game of incomplete information, $\mathbf{\Gamma} = \langle N, \mathbf{T}, \mathbf{p}, \mathbf{A}, \mathbf{\Delta}, \mathbf{U}, \mathbf{P}\rangle$. Let $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, ..., \boldsymbol{\sigma}_n)$ be a profile of strategy vectors, one for each player, where $\boldsymbol{\sigma}_j = \left(\sigma_j(.,t_{j1}), \sigma_j(.,t_{j2}), ..., \sigma_j\left(.,t_{jn_j}\right)\right)$ is a vector of strategies for player j, one for each type of player j. Let $\mathbf{P}_{it_i,jt_j}(\sigma_i(.,t_i)) = \sigma_j^e(.,t_j), j \neq i$. Let $\boldsymbol{\sigma}_j^e(\sigma_i(.,t_i)) = \left(\sigma_j^e(.,t_{j1}), \sigma_j^e(.,t_{j2}), ..., \sigma_j^e\left(.,t_{jn_j}\right)\right), j \neq i$. Let $\boldsymbol{\sigma}_{-i}^e(\sigma_i(.,t_i)) = \left(\boldsymbol{\sigma}_1^e(.), \boldsymbol{\sigma}_2^e(.), ..., \boldsymbol{\sigma}_{i-1}^e(.), \boldsymbol{\sigma}_{i+1}^e(.), ..., \boldsymbol{\sigma}_n^e(.)\right.$ Then:*
(a) $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, ..., \boldsymbol{\sigma}_n)$ is an evidential equilibrium if $\sigma_i(.,t_i)$ is an optimal strategy for each type, t_i, of player i and for each player $i \in N$ (recall Definitions 13, 21 26).
(b) $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, ..., \boldsymbol{\sigma}_n)$ is mutually consistent if $\boldsymbol{\sigma}_{-i}^e(\sigma_i(.,t_i)) = \boldsymbol{\sigma}_{-i}(\sigma_i(.,t_i))$ for each type, t_i, of player i and for each player $i \in N$.
(c) $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, ..., \boldsymbol{\sigma}_n)$ is a consistent evidential equilibrium if it satisfies both (a) and (b), above.

Proposition 4 *: (a) Let the profile of strategy vectors $\boldsymbol{\sigma}^*$ be a Nash equilibrium in the static game of incomplete information, $\mathbf{G} = \langle N, \mathbf{T}, \mathbf{p}, \mathbf{A}, \mathbf{\Delta}, \mathbf{U}\rangle$. Let*

$$\boldsymbol{\sigma}_j = \left(\sigma_j(.,t_{j1}), \sigma_j(.,t_{j2}), ..., \sigma_j\left(.,t_{jn_j}\right)\right), \ j \in N.$$

Consider the (constant) social projection functions: $\mathbf{P}_{it_i,jt_j}(\sigma_i(.,t_i)) = \sigma_j^*(.,t_j)$, $t_i \in T_i$, $i \in N$, $j \in N - \{i\}$, $t_i \in T_i$. Then $\boldsymbol{\sigma}^*$, is a consistent evidential equilibrium in the evidential game $\mathbf{\Gamma} = \langle \mathbf{G}, \mathbf{P}\rangle$. Furthermore, Γ is a causal game.
(b) Let the profile of strategy vectors $\boldsymbol{\sigma}^*$ be an evidential equilibrium in the evidential game $\Gamma = \langle N, \mathbf{T}, \mathbf{p}, \mathbf{A}, \mathbf{\Delta}, \mathbf{U}, \mathbf{P}\rangle$, where $\mathbf{P}$ is the family of constant social projection functions $\mathbf{P}_{it_i,jt_j}(\sigma_i(.,t_i)) = \sigma_j^*(.,t_j)$, $t_i \in T_i$, $i \in N$, $j \in N - \{i\}$, $t_i \in T_i$ (hence, $\boldsymbol{\sigma}^*$ is a consistent

evidential equilibrium and Γ is a causal game). Then $\boldsymbol{\sigma}^*$ is a Nash equilibrium in the static game of incomplete information $\mathbf{G} = \langle N, \mathbf{T}, \mathbf{p}, \mathbf{A}, \mathbf{\Delta}, \mathbf{U} \rangle$.

Proof of Proposition 4: Similar to that of Proposition 1. ■.

6.3.1 An example

We can now use the machinery of evidential games of incomplete information to formalize the discussion in Example 5. Let player 1 be the row player in Table 6, and player 2 the column player. Player 1 has only one type, $T_1 = \{t_{11}\}$ while player 2 has two types, $T_2 = \{t_{21}, t_{22}\}$. In Example 5, we assume that: (1) Player 1 assigns a probability p that player 2 is of type t_{21}; a type who always defects (2) Player 1 assigns the identity social projection function to type t_{22}. Thus, the social projection function of player 1 is given by

$$\begin{cases} P_{11,21}(C|C) = 0, P_{11,21}(D|C) = 1; P_{11,21}(D|D) = 1, P_{11,21}(C|D) = 0, \\ P_{11,22}(C|C) = 1, P_{11,22}(D|C) = 0; P_{11,22}(C|D) = 0, P_{11,22}(D|D) = 1, \end{cases} \tag{31}$$

thus, for example, $P_{11,21}(D|C) = 1$ says that the first (and only) type of player 1 anticipates that if he plays C then the first type of player 2 will defect. In the first row of (31), Player 1 uses causal reasoning towards type t_{21} of player 2 and in the second row he uses the identity social projection function towards type t_{22} of player 2. For illustrative purposes, we assume that type t_{21} of player 2 uses causal reasoning towards player 1, while type t_{22} of player 2 uses evidential reasoning towards player 1. This gives rise to the following social projection functions for the two types of player 2.

$$\begin{cases} P_{21,11}(C|C) = 0, P_{21,11}(D|C) = 1; P_{21,11}(D|D) = 1, P_{21,11}(C|D) = 0, \\ P_{22,11}(C|C) = 1, P_{22,11}(D|C) = 0; P_{22,11}(C|D) = 0, P_{22,11}(D|D) = 1. \end{cases} \tag{32}$$

Given (31), for player 1, cooperation is better than defection if $p \times 0 + (1-p)2 \geq 1$ or $p \leq \frac{1}{2}$. Thus, if player 1 assigns more than an even probability of player 2 being like-minded then he will cooperate. Using (32), since type t_{21} of player 2 uses causal reasoning to form beliefs about player 1, he always finds it optimal to defect. On the other hand, since type t_{22} of player 2 believes that player 1 is like-minded, he always finds it optimal to cooperate.

Suppose that $p \leq \frac{1}{2}$. Then player 1 and type t_{22} of player 2 both find it optimal to cooperate. Should they be matched, their initial beliefs are fulfilled. If however, player 1 and type t_{21} of player 2 are matched then player 1 will get a payoff of 0 and player 2 will unexpectedly get a higher payoff of 3; initial beliefs are not fulfilled in this case.

Suppose that $p > \frac{1}{2}$. In this case, if player 1 and type t_{21} of player 2 are matched then both defect and initial beliefs are fulfilled. If, however player 1 and type t_{22} of player 2 are matched then player 1 unexpectedly gets a higher payoff of 3 and initial beliefs are not upheld. Recall that since the social projection functions of players are not known to others, other players cannot exploit their social projection functions for strategic advantage.

7 Conclusions

In static games of complete or incomplete information, players are uncertain about which actions the other players will take. Aumann and Brandenburger (1995) gave epistemic conditions under which the play of a game would result in a Nash equilibrium. A very large number of experimental subjects do not play a Nash equilibrium in well known games such as prisoners' dilemma, voting games, public goods games and oligopoly games. A violation of Nash equilibrium is also a violation of the epistemic conditions that underlie it. It would seem that this constitutes strong grounds for game theory to be open to alternative equilibrium concepts in static games.

A great deal of evidence suggests that in resolving uncertainty about what other like-minded players will do, players assign *diagnostic significance* to their own actions. Such reasoning is described as *evidential reasoning* (ER) but it is disallowed by the epistemic conditions in Aumann-Brandenburger (1995). Often players use ER without being aware of using it. Other evidence suggests that it is an automatic response. In other words, humans might be hard-wired to use ER (possibly for evolutionary reasons). Finally, players using ER do not believe that their actions cause others to take any particular actions.

The aim of our paper is to explore the significance of ER for the class of static games of complete and incomplete information. We define evidential games (EG) in which some players use ER. We also propose the relevant solution concepts for such games: *Evidential equilibrium* (EE) and *consistent evidential equilibrium* (CEE). In the latter (but not necessarily the former) beliefs turn out to be correct in equilibrium.

We give applications of EE in several common games, in particular, the prisoners' dilemma and oligopoly games. In each case, ER produces a greater degree of cooperation relative to a NE. If the cooperative outcome is associated with a higher payoff that players prefer, then they often take their own preference for cooperation as being of diagnostic significance about the likely cooperation of others. The evidence shows that the cooperative outcome in prisoners' dilemma is played about 50% of the times despite the strategy *defect* strictly dominating the strategy *cooperate*. Similarly a great deal of evidence shows that the outcome in oligopoly games, particularly when players have uncertainty about others, is not a Cournot Nash equilibrium and the collusive outcome is not uncommon. In each of these cases, players do not require an infinite horizon (or a finite horizon with the conventional degree of irrationality) or other regarding preferences in order to cooperate. These factors may well be very important but ER provides an alternative foundation for cooperative behavior.

It is very likely that there could be a mixture of players: Some use ER while others use the conventional reasoning in game theory (following standard usage in psychology, we call it causal reasoning). Indeed in our analysis, we allow for such a mixture. Future research, both empirical and theoretical, could fruitfully explore this idea and work on the estimation of such mixtures.

Our framework can naturally be extended to dynamic games but we lack a body of evidence that could underpin such an extension. Hence, we leave such developments for future research as more evidence accumulates.

Acknowledgements: We are very grateful to Chris Wallace and Andrew Colman for their comments on the paper.

References

Acevedo, M. and Krueger, J. I. (2005) Evidential reasoning in the prisoners' dilemma game, American Journal of Psychology, 118: 431-57.

Aumann, R. J. and Brandenburger, A. (1995). Epistemic Conditions for Nash Equilibrium, *Econometrica*, 63:1161–1180.

Bosch-Domènech, A. Vriend, N. J. (2003): "Imitation of successful behaviour in Cournot markets," Economic Journal, 113: 495–524.

Camerer, Colin F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton, NJ: Princeton University Press.

Colman, Andrew, Briony Pulford, David Omtzigt and Ali al-Nowaihi. 2010. Learning to co-operate without awareness in multiplayer minimal social situations, *Cognitive Psychology*, 61, pp201-227. Available online doi:10.1016/j.cogpsych.2010.05.003.

Crawford, Vincent and Nagore Iriberri, 2007. "Level-k Auctions: Can a Non-quilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?," Econometrica, Econometric Society, 75: 1721-1770.

Cooper, R., DeJong, D. V., Forsythe, R., and Ross, T. W., 1996. "Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games," Games and Economic Behavior, 12: 187-218.

Darai, D., and Grätz, S. (2010) Golden Balls: A Prisoner's Dilemma Experiment, University of Zurich Working Paper 1006.

Dawes, Robyn M & Thaler, Richard H, (1988). Anomalies: Cooperation, *Journal of Economic Perspectives*, 2: 187-97.

de Sousa, J., Hollard, G. and Terracol, A. (2012) Do non-strategic players really exist? Evidence from experimental games involving step reasoning, mimeo University of Paris.

Delavande, A. and Manski, C. F. (2012) Candidate preferences and expectations of election outcomes. PNAS, 109: 3711-3715.

Duersch, P., Kolb, A., Oechssler, J., & Schipper, B. (2010). Rage against the machines: how subjects play against learning algorithms. Economic Theory, 43, 407–430.

Fehr, E., and Gächter, S., 2000. Cooperation and punishment in public goods experiments. American Economic Review 90: 980–994.

Fehr, E., and Schmidt, K. M., (2010) On inequity aversion: A reply to Binmore and Shaked," Journal of Economic Behavior & Organization, 73: 101-108.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. Quarterly Journal of Economics 114, 817-868.

Fischbacher, U., Gachter, S., Fehr, E., (2001). "Are people conditionally cooperative? Evidence from a public goods experiment," Economics Letters, 71: 397-404.

Forsythe, R., F. Nelson, GR. Neumann and J. Wright (1992) "Anatomy of an Experimental Political Stock Market." American Economic Review, 82:1142-1161.

Fouraker, L., and Siegel, S., 1963. Bargaining Behavior. McGraw-Hill, New York.

Gächter, S. and Thöni, C. (2005). "Social Learning and Voluntary Cooperation Among Like-Minded People," Journal of the European Economic Association, 3: 303-314.

Gintis, H. (2012) Common Knowledge of Rationality is Self-Contradictory, mimeo University of Massachusetts.

Grafstein, R. (1991) An evidential decision theory of voter turnout, American Journal of Political Science, 35: 989-1010.

Halpern, J. (1986). Reasoning over knowledge: An overview, in J. Halpern (ed.), *Theoretical aspects of reasoning about knowledge*. Morgan Kaufmann.

Howard, J. V. (1988) Cooperation in the prisoners' dilemma, Theory and Decision, 24: 203-213.

Huck, S., Normann, H. T., & Oechssler, J. (1999). Learning in Cournot oligopoly—an experiment. *Economic Journal*, 109, 80–95.

Ivanov, A., Levin, D. and Niederle, M. (2010), Can Relaxation of Beliefs Rationalize the Winner's Curse?: An Experimental Study. Econometrica, 78: 1435–1452.

Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics, American Economic Review, 93: 1449-1475.

Kahneman, D., Slovic, P., and Tversky, A. (Eds.). (1982), Judgement Under Uncertainty: Heuristics and Biases, Cambridge University Press, Cambridge.

Kahneman, D. and Tversky, A. (2000) Choices, Values and Frames. Cambridge University Press: Cambridge.

Khadjavi, M., and Lange, A. (2013) "Prisoners and their Dilemma", Journal of Economic Behavior and Organization, 92: 163-75.

Koudenburg, N., Postmes, T., Gordijn, E. H. (2011). If they were to vote they would vote for us. *Pscyhological Science*. 22: 1506-1510.

Kreps, D., Milgrom, P., Roberts, J., and Wilson, R. (1982). "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma," Journal of Economic Theory, 17: 245–252.

Krueger, J. I. (2007) From social projection to social behavior, *European Review of Social Psychology*, 18: 1-35.

Krueger, J. I. and Acevedo, M. (2008) A game-theoretic view of voting, *Journal of Social Issues*, 64: 467-85.

Lenton, A. P., Bryan, A., Hastie, R., and Fischer, O. (2007). We want the same thing: Projection in judgements of sexual intent. *Personality and Social Psychology Bulletin*. 33: 975-988.

Lewis, D. K. (1979). Prisoners' dilemma is a Newcomb problem. *Philosophy and Public Affairs*, 8: 235-240.

Nozick, R. (1993). *The nature of rationality*. Princeton, NJ: Princeton University Press.

Polak, B. (1998) "Epistemic Conditions for Nash Equilibrium, and Common Knowledge of Rationality", Econometrica, 67: 673-676.

Quattrone, G. A., and Tversky, A. (1984). Causal versus diagnostic contingencies: On self deception and the voter's illusion. *Journal of Personality and Social Psychology*, 46: 237-248.

Rapoport, A. (1988). Experiments with N-Person Social Traps I: Prisoners' Dilemma, Weak Prisoners' Dilemma, Volunteers' Dilemma, and Largest Number, *Journal of conflict resolution*, 32: 457-472.

Rapoport, A. 1966, Two-Person Game Theory, Michigan University Press: Ann Arbor.

Rassenti, S., Reynolds, S. S., Smith, V. L. and Szidarovszky, F., (2000) Adaptation And Convergence Of Behavior In Repeated Experimental Cournot Games. Journal of Economic Behavior & Organization, 41: 117-46.

Requate, T. and Waichman, I. (2011). A profit table or a profit calculator? A note on the design of Cournot oligopoly experiments, *Experimental Economics*, 14: 36-46.

Riketta, M. & Sacramento, C. A. (2008). They Cooperate With Us, So They Are Like Me: Perceived Intergroup Relationships Moderates Projection From Self to Outgroups. *Group Processes and Intergroup Relations*. 11: 115-31.

Robbins, J. M., and Krueger, J. I. (2005). Social projection to ingroups and outgroups: A review and meta-analysis. *Personality and Social Psychology Review*, 9: 32-47.

Ross, L., Greene, D., and House, P., (1977). The 'False Consensus Effect': An Egocentric Bias in Social Perception and Attribution Processes; *Journal of Experimental Social Psychology*, 13, 279-301.

Van Boven, L., and Loewenstein, G. (2005). Cross-situational projection. In M. Alicke, J. Krueger, & D. Dunning (Eds.), *The Self in Social Judgment*, pp. 43–64, Psychology Press: New York.

Waichman, I., Requate, T., and Siang, C. K. (2010). A Cournot experiment with managers and students: evidence from Germany and Malaysia. *The BE Journal of Economic Analysis and Policy* (Topics) 10, Article 30.

Zhong, Chen-Bo., Loewenstein, J. and Murnighan, J. K. (2007). Speaking the Same Language: The Cooperative Effects of Labeling in the Prisoner's Dilemma, *Journal of Conflict Resolution*, 51: 436-56.