University *of* Leicester

DEPARTMENT OF ECONOMICS

Evidential Equilibria in Static Games

Sanjit Dhami, University of Leicester, UK

Ali al-Nowaihi, University of Leicester, UK

Working Paper No. 12/15
July 2012

Evidential equilibria in static games

Sanjit Dhami[*] **Ali al-Nowaihi**[†]

16 July 2012

Abstract

Under uncertainty about what others will do, 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 others. ER successfully explains the evidence from many important games. We provide a formal theoretical framework for discussing ER by proposing *evidential games* and the relevant solution concept-*evidential equilibrium* (EE). We derive the relation between a Nash equilibrium and an EE. We apply EE to several common games including the prisoners' dilemma and oligopoly games.

Keywords: Evidential and causal reasoning; evidential games; social projection functions; ingroups and outgroups; evidential equilibria and consistent evidential equilibria; Nash equilibria, common knowledge and epistemic foundations.

[*]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

Psychologists have found considerable evidence for a type of reasoning that they call *evidential reasoning* (ER), also sometimes referred to as *social projection.* When players who use ER are uncertain of the actions of other players they assign *diagnostic significance* to their own actions in forming beliefs about the actions of others. In other words, players take their own action as evidence of how "similar" players would behave. 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".

Following the psychological literature, we refer to the standard reasoning used in economics as *causal reasoning.* A decision maker who uses causal reasoning assigns no diagnostic value to his own actions, unlike ER. Hence, a causal decision maker does not alter his beliefs about the distribution of the types of others from the actions that *he himself* takes. But, of course, a causal decision maker may alter his beliefs (using Bayes' law, for example) about the distribution of the types of others from the actions that *they* take.

An illustrative example may help. Consider the standard *one-shot prisoners' dilemma* game. A player who uses causal reasoning, and perceives the other player to be 'like minded', will reason as follows "To defect is a strictly dominant strategy for me, so I will defect". On the other hand, a player who uses evidential reasoning 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. I expect that my rival, who is thinking like me, will also cooperate."

Empirical evidence (section 2) shows that there is far greater cooperation in the prisoners' dilemma experiments relative to the predictions under causal reasoning. Furthermore, players who choose to play the action cooperate (defect) estimate with high probability that others will cooperate (defect). One natural implication, therefore, is that welfare under ER may be higher than under causal reasoning.[1] We show in section 2 that the evidence is suggestive of similar results from many other contexts, such as public goods, voting, false consensus effect, projection bias etc. In each case, decision makers who take a particular action assign a much higher probability that other like-minded people will also

[1] This observation suggests that ER could be the product of human evolution. Decision makers who use ER may achieve a better cooperative outcome relative to those who use causal reasoning. Hence, ER may well turn out to be an evolutionary stable strategy in some well defined evolutionary game. The evidence suggests (see below) that ER is an automatic response and cognitive effort is required to suspend it. This is suggestive of the view that humans may be hard-wired to use ER.

take the same action.

Economists accustomed to traditional notions of rationality might find evidential reasoning to be less than fully rational. However, the relevant issue is the evidence on human behavior. Our own reading of the evidence supports the importance of evidential reasoning. It may well be that there is a mixture in the population, some who use ER and others who use causal reasoning. Our framework allows for such a mixture.

Section 2 of this paper, *Evidential Reasoning*, gives a more detailed, though still informal, discussion of evidential reasoning including further examples.

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 is optimizing 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. Our formulation of evidential reasoning yields causal reasoning, the dominant reasoning assumed 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 section 3.4.

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 (section 4) and oligopoly games (section 5). We review the empirical evidence on these games. We conclude that the evidence from these games is more supportive of evidential reasoning than causal reasoning.

Section 6 concludes.

2. Evidential Reasoning

When players are uncertain about what their opponents will do, the evidence suggests that players take their own actions as diagnostic of what like-minded players might do. If, however, they believe that the other players are not like-minded then they might not assign diagnostic significance to their actions (this issue is discussed in greater length in section 3.2 below). We now offer further discussion and some concrete contexts in which evidential reasoning has been studied. In this section we shall use the terms *evidential reasoning* and *causal reasoning* as informally defined in the introduction. Formal definitions are given in Section 3, below.

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 can explain cooperative behavior even if individuals behave in a self interested manner. This is in contrast with other models of behavior where cooperation is explained by assuming that individuals have social preferences (or other-regarding preferences). We now give several examples of evidential reasoning.

2.1. False consensus and evidential reasoning about others

Ross et al. (1977) asked experimental subjects if they would walk around a university campus wearing a sandwich board that said "REPENT". Those who agreed also estimated that 63.5% of their peers would do so too, 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 be *consistent beliefs*. This evidence is consistent with subjects using evidential reasoning to impute diagnostic value to their own 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*.[2]

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

The static prisoners' dilemma is a well known game in social sciences in which each of the two players can either cooperate or defect. The payoffs from each action to the row and the column player are shown in the following payoff matrix. In each cell the first payoff is to the row player.

	Cooperate	Defect
Cooperate	2,2	0,3
Defect	3,0	1,1

Defection is a strictly dominant strategy- irrespective of what the other player does it gives a strictly higher payoff. Hence, a player using causal reasoning should defect *no matter what probability* $p \in [0, 1]$ *he assigns to the other player of cooperating.* The unique

[2] Mullen (1985) surveyed 115 studies of the false consensus effect. He defined the term false consensus thus: "False consensus refers to an egocentric bias that occurs when people estimate consensus for their own behaviors. Specifically, the false consensus hypothesis holds that people who engage in a given behavior will estimate that behavior to be more common than it is estimated to be by people who engage in alternative behaviors".

prediction of standard game theory (and the unique Nash equilibrium of the game) is that each player defects and achieves a payoff of 1.

In contrast, experimental evidence indicates high cooperation rates. In a study of the prisoners' dilemma game based on high stakes outcomes from a British TV show called Goldenballs, Darai and Grätz (2010) find unilateral cooperation rates of 55% for stakes above £500 and cooperation rates of 74% for stakes below this level. Rapoport (1988) finds cooperation rates of 50% in the prisoners dilemma game. 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"). When purely generic labels are used (such as C and D) then the cooperation rates are about 50%.[3] Even if players correctly figure out that other players will cooperate with high probability, say $p = 0.50 - 0.74$ (as the figures above suggest), it is still not optimal for a player who uses causal reasoning to cooperate. Hence, within the domain of causal reasoning, the puzzle remains as to why we observe any cooperation at all in the prisoners' dilemma game.

Lewis (1979) used evidential reasoning to explain these unexpected levels of cooperation in the one-shot prisoners' dilemma game. The payoff from mutual cooperation is better than mutual defection. If players use evidential reasoning, they might 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 believe that the probability of other players cooperating is between 0.6 and 0.7. Similarly, players who defect believe that other players will defect with probabilities between 0.6 to 0.7; see Krueger (2007).[4] These issues are taken up in more detail in section 4 below.

Clearly, not everyone chooses to cooperate in the prisoner's dilemma game-a sizeable minority defects. A plausible explanation is that there could be a mixture of players in the population- a majority who use evidential reasoning and a sizeable minority who use standard causal reasoning. Our formal model will allow for such mixtures of players. This comment also applies to the other examples that we offer in this section.

[3] 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).

[4] The experimental evidence from Kay and Ross (2003) shows that individuals cooperate more if the prisoners' dilemma game is framed as a cooperative game rather than as a competitive game. This piece of evidence is inconsistent with the *frame-invariance* assumption of classical economic theory. However, it supports evidential reasoning. When framed as a cooperative game, players assign even higher diagnostic value to their own preference for cooperation over defection.

2.3. Why do people voluntarily contribute towards a public good?

In public good games a set of individuals contribute simultaneously towards a *public good* that gives utility to all individuals.[5]

A typical experiment proceeds as follows. Suppose that there is a group of 6 subjects. Each subject tries to maximize his/her monetary payoff and each is endowed with £20. All experimental subjects $i = 1, ..., 6$ simultaneously contribute an amount $x_i \in [0, 20]$ towards the public good. The total amount of the public good is $G = \sum_{i=1}^{6} x_i$. Since the public good is non-rival and non-excludable, the experimenter then gives each consumer an "identical" amount rG, $0 < r < 1$, to capture benefits from the public good. Suppose $r = 0.4$. The 'first best' (which maximizes the joint payoffs of all players) is for each individual to contribute $x_i = 20$ for a total monetary payoff of £48. However, conditional on everyone cooperating fully, if any individual *free rides* (i.e., contributes $x_i = 0$) then his/her payoff is $20 + 0.4(100) = 60 > 48$. Thus, cooperation cannot be sustained as a Nash equilibrium and everyone would prefer to free ride ($x_i = 0$, $i = 1, 2, ..., 6$) for a monetary payoff of £20, which is just the original endowment. This is the unique prediction under causal reasoning (and the unique Nash equilibrium of the game).

In experiments, in early rounds of public good games, individuals contribute quite high levels that are between a half to three-quarters of their maximum possible contributions.[6] This contradicts the unique prediction under causal reasoning. Evidential reasoning provides a possible explanation for the observation of cooperation in the early rounds of public good games. Many players take their own desire for mutual contribution (for a outcome of £48 rather than £20 in the above example) in the first round as diagnostic of the desire of other players to contribute, hence, they contribute.[7] Contributions drop off in subsequent rounds unless ex-post punishment of non-cooperators by cooperators is allowed. This is on account of negative reciprocity; see Fehr and Gächter (2000). Thus, in conjunction, evidential reasoning and negative reciprocity give a good account of the behavior in public good games.

The role of evidential reasoning in the early rounds of the public good games, is supported by the evidence in Gächter and Thöni (2005). They investigate whether cooperation in public good games is higher among 'like-minded' people. In order to separate the sub-

[5] Two classic examples of public goods are a lighthouse and national defence. A public good is *non-rival* (i.e., one person's consumption does not reduce consumption of others and *non-excludable* (i.e., it is not possible to exclude users of the good).

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

[7] For the dynamic version of the public good game (where the contributions of each player are revealed at the end of each round) we conjecture the following. Evidential reasoning (for the first few rounds) in conjunction with negative reciprocity (for subsquent rounds) gives a better description of the evidence from public good games. But at the moment we lack empirical evidence on the type, extent and nature of evidential reasoning in repeated games.

Year	Presidential candidates	% intending to vote for the Democrats who expect Democrats to win	% intending to vote for the Republicans who expect Democrats to win
1988	Dukakis/Bush	51.7	5.8 (94.2)
1984	Mondale/Reagan	28.8	1.0 (99.0)
1980	Carter/Reagan	87.0	19.6 (80.4)
1976	Carter/Ford	84.2	19.6 (80.4)
1972	McGovern/Nixon	24.7	0.4 (99.6)
1968	Humphrey/Nixon	62.5	4.6 (95.4)
1964	Johnson/Goldwater	98.6	69.5 (30.5)
1960	Kennedy/Nixon	78.4	15.8 (84.2)
1956	Stevenson/Eisenhower	54.6	2.4 (97.6)
1952	Stevenson/Eisenhower	81.4	14.1 (85.9)

Figure 2.1: Source: Forsythe et al. (1992)

jects into like-minded people they initially run a single-round 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. Over the next 10 rounds, contributions are much higher and free riding much lower among groups whose members contributed the most in round one. In particular, the contributions of the TOP group approach the first best level in several rounds.[8]

2.4. Why do people vote and how do they form beliefs?

Quattrone and Tversky (1984, 1988) used evidential reasoning to explain the *voting paradox*. Under causal reasoning, given that any one voter is most unlikely to be pivotal, it is in nobody's interest to vote.[9] But then why do so many people vote?

Under evidential reasoning one takes one's own actions as diagnostic of what other like-minded people are likely to do. Hence, a voter using evidential reasoning could reason 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 has a better chance of winning. So I vote if I wish my party to win, otherwise I do not." It is important to note that one's action to vote does not *cause* others to vote- it

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

[9] 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).

only has diagnostic significance about what other, like minded, voters are likely to do.

A corollary to the argument given above is that if I use evidential reasoning and decide to vote for a particular political party then I also think that all other like-minded people will vote for that party. I might then assign too high a probability for my preferred party to win the election because not all like-minded people might actually vote as I believe they would. 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.

Figure 2.1 is based on survey data from successive US Presidential elections. Voters who intend to vote Democrats typically assign high probabilities to the Democrat candidate winning. In contrast, voters who intend to vote Republican (last column in the Figure) assign low probabilities to a Democrat win and high probabilities to a Republican win (the latter are shown in brackets in the last column). Thus, voters seem to take their own actions as diagnostic of what other like minded people will do, supporting the evidential reasoning explanation.

Delavande and Manski (2011) 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 the same estimates of which party will win. But since voters might be using evidential reasoning, they assign too high a probability to their preferred party winning the election.[10] A natural implication is that *beliefs are not mutually consistent.* If voters were using causal reasoning and the same information were available to all voters then the percentages in column 3 and the bracketed percentages in column 4 of Figure 2.1 should add up to a 100% (in contrast in the last row of Figure 2.1, for the 1952 election, these figures add up to 167.3%!). However, there is no inconsistency in the figures if one allows for evidential reasoning because there is no presumption that beliefs be mutually consistent under such a form of reasoning.

2.5. 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

[10] Their findings are invariant with respect to males/females, whites/non-whites, educated/non-educated etc. Hence, there is a strong possibility that evidential reasoning is hard-wired in humans.

only if he receives confirmation that his message has reached W, and so on. Under causal reasoning, neither W nor B will attack. However, under evidential reasoning, W and B will both attack because each uses his own reasoning as evidence that the other will attack (and, maybe, a finite number of messages is sufficient to enforce this psychological mode of reasoning).[11] 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, have practical utility.

2.6. 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 with or cue value for the lost hiker. In each case, they assign *diagnostic significance* to their own state in inferring the state of the lost hiker.

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.

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

2.7. Projection bias and multiple selves

An example of evidential reasoning is the well established phenomenon of *projection-bias*; see Loewenstein et al. (2003). In a rich literature in psychology, the individual is viewed as a succession of multiple selves, one for each time period. Individuals who have projection-bias overly use the preferences of their current self to predict the preferences of their future selves. For instance, shoppers who shopped for groceries for the next week purchased more groceries when hungry during shopping than when satiated. In each case, individuals seem

[11] To elaborate, suppose that both W and B believe that they are likeminded and use evidential reasoning. If any of them does not wish to attack then he takes it is evidence that the other will not attack, so both lose. If any of them wishes to attack then he takes this as evidence that the other will also attack, in which case N will be defeated- a better outcome for W and B. Thus, both attack, solving the coordination problem.

to be assigning diagnostic significance to the preferences of their current selves to impute preferences to their future selves.

2.8. Calvinism and the development of capitalism

This example is considered in 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, an explanation that is popular in sociology; see for example, Nozick (1993).[12]

3. Evidential Games

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.[13] $\mathbf{A} = \prod_{i=1}^{n} A_i$ 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 $\boldsymbol{\Delta} = \prod_{i=1}^{n} \Delta_i$ is the set of all possible profiles of strategies. A particular profile of strategies of other players is denoted by $\boldsymbol{\sigma}_{-i} = (\sigma_1, ..., \sigma_{i-1}, \sigma_{i+1}, ..., \sigma_n) \in \boldsymbol{\Delta}_{-i} = \prod_{j \in N - \{i\}} \Delta_j$.

The payoff of player i is $\pi_i : \Delta \to \mathbb{R}$ 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.[14] In an experimental setup, this 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,

---

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

[13] 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.

[14] *Common Knowledge* was first informally defined by Lewis (1969) and formally by Aumann (1976). The latter also proved the famous theorem on "agreeing to disagree".

chooses his strategy, σ_i, he does not know the strategies, $\boldsymbol{\sigma}_{-i}$, chosen by the other players. This is the sense in which this is a *static 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", $\mathbf{P}$; this is undertaken in subsection 3.1, below.

3.1. 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. We will call such a function a *social projection function*.[15]

Definition 2 *(Social projection functions):*
(a) A social projection function for player i (SPF for short), is a mapping $\mathbf{P}_i : \Delta_i \rightarrow \boldsymbol{\Delta}_{-i}$, that assigns to each strategy, $\sigma_i \in \Delta_i$, for player i, the $(n-1)$ vector of strategies, $\mathbf{P}_i\left(.|\sigma_i\right)$, for the other players. We describe such a player as using evidential reasoning.
(b) We write, $P_{ij}\left(.|\sigma_i\right)$, for the j-th component $(j \neq i)$ of $\mathbf{P}_i\left(.|\sigma_i\right)$. We write $P_{ij}\left(a_j|\sigma_i\right)$ for the probability player i assigns to player j playing $a_j \in A_j$, conditional on player i playing $\sigma_i \in \Delta_i$. Hence, $P_{ij}\left(a_j|\sigma_i\right) \geq 0$ and $\sum_{a_j \in A_j} P_{ij}\left(a_j|\sigma_i\right) = 1$.
(c) Player i regards player j $(j \neq i)$ as an outgroup member if $P_{ij}\left(a_j|\sigma_i\right)$ is independent of σ_i, i.e., if $P_{ij}\left(a_j|\sigma_i\right) = P_{ij}\left(a_j|\sigma_i'\right)$ 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.
(d) 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.
(e) 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.
(f) We say that player i uses causal reasoning if $\mathbf{P}_i\left(.|\sigma_i\right)$ is independent of σ_i, i.e., if $\mathbf{P}_i\left(.|\sigma_i\right) = \mathbf{P}_i\left(.|\sigma_i'\right)$ for all $\sigma_i, \sigma_i' \in \Delta_i$.

[15] 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.

We now offer some remarks that discuss various aspects of Definition 2 and the relation between parts (c) and (f).

Remark 1 : *In a static game, 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 2 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. On the other hand, if player i uses causal reasoning (as in classical game theory) then he assigns no diagnostic significance to his own strategy, σ_i, in inferring the strategies, $\boldsymbol{\sigma}_{-i}$, followed by the other players.*[16] *Thus, under causal reasoning, $\mathbf{P}_i(.|\sigma_i)$ remains fixed as σ_i changes.*

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 $P_{ij}(a_j|\sigma_i) = \sigma_{ij}(a_j)$ (which, of course, depends on i but is independent *of σ_i). Hence, in this case,*

$$P_{ij}(a_i, a_j|\sigma_i) = \sigma_i(a_i) P_{ij}(a_j|\sigma_i) = \sigma_i(a_i) \sigma_{ij}(a_j).$$

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. This explains relation between parts (c) and (f) of Definition 2.

In Section 2, in the prisoner's dilemma game (section 2.2), or the public goods game (section 2.3) or in the other experiments reported there, players taking a particular action appeared to believe that like minded players (part of the ingroup) will also take an identical action. This leads to the idea of an identity social projection function in Definition 3.

[16] 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.

Definition 3 *(Identity social projection function): Let $M \subseteq N$ be a subset of players. Suppose 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 $P_{ij}(a|\sigma_i) = \sigma_i(a)$ 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.*

When all players in the ingroup use the identity social projection function, we call this a perfect ingroup in Definition 4, below.

Definition 4 *(Perfect ingroups): Let $M \subseteq N$ be a subset of players. Suppose 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.*

We now have the machinery to define an evidential game and a causal game.

Definition 5 *(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 $\Gamma = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$. We call Γ an evidential game and 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 *: Note the following:*
(a) According to Definition 2, causal reasoning is a special case of evidential reasoning and according to Definition 5 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 1 *: Consider the two player 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 $A_1 = A_2 = \{H, T\}$. Let p be the probability with which player 1 plays H (hence, player 1 plays T with probability $1 - p$) and let q be the probability with which player 2 plays H (hence, player 2 plays T with probability $1 - q$). 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,

$$P_{12}(H \mid p) = p,\ P_{21}(H \mid q) = 0.5 \text{ for all } p, q \in [0, 1]\ . \tag{3.1}$$

According to (3.1), player 1 believes that if he (player 1) plays H with probability p, then so will player 2. Hence, player 1 has an identity social projection function. Player 2, however, 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 (in fact, player 2 exhibits causal reasoning). 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; see Definition 4). By contrast, if in (3.1) we had $P_{12}(H \mid p) = 0.3$ (say), then both players would exhibit causal reasoning; and this example would become a causal game.

3.2. 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" (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 2c,d,e).

However, recent evidence suggests a more nuanced view (and this is, again, consistent we 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 only the set of voters forms an ingroup (Definition 2c,d).

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 2.6). They find that an ingroup member may have a harmonious (or cooperative) relation with other ingroup members. On the other hand, they also find that an ingroup member may be in competition (or conflict) with other ingroup members. In the latter case, an ingroup member may believe that the actions of others are in contrast to his own actions (the *contrast effect*). Subsection 5.2 will give further examples of these.

3.3. Equilibria

Definition 6 *(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 5), is one that maximizes the payoff function,*

$\pi_i\left(\sigma_i, \mathbf{P}_i\left(.|\sigma_i\right)\right)$, *of player* i.

Definition 7 *(Evidential equilibria): The strategy profile* $\boldsymbol{\sigma}^* = \left(\sigma_1^*, \sigma_2^*, ..., \sigma_n^*\right) \in \boldsymbol{\Delta}$ *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 6).*

Note that Definition 7 only requires that a strategy for a player be optimal given *his beliefs*. But, of course, beliefs may not turn out to be correct (see for instance, section 2.4 above). 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. 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.

Definition 8 *(Mutually consistent strategies): A strategy profile* $\boldsymbol{\sigma}^* = \left(\sigma_1^*, \sigma_2^*, ..., \sigma_n^*\right) \in \boldsymbol{\Delta}$ *of the evidential game* $\boldsymbol{\Gamma} = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$ *(Definition 5) is a mutually consistent vector of strategies if* $\mathbf{P}_i\left(.|\sigma_i^*\right) = \boldsymbol{\sigma}_{-i}^*$ *for all* $i \in N$*, i.e., if* $P_{ij}\left(a_j|\sigma_i^*\right) = \sigma_j^*\left(a_j\right)$*, for all* $i, j \in N$*,* $i \neq j$*, and all* $a_j \in A_j$*.*

In other words, a strategy profile $\boldsymbol{\sigma}^* = \left(\sigma_1^*, \sigma_2^*, ..., \sigma_n^*\right)$ 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 $P_{ij}\left(a_j|\sigma_i^*\right)$ that player i 'thinks' player j will play action a_j (given σ_i^*) is equal to the probability $\sigma_j^*\left(a_j\right)$ with which player j 'actually' plays a_j.

Definition 9 *(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 7 and 8).*

3.4. Nash equilibria and consistent evidential equilibria

As one might expect, there is a natural correspondence between Nash equilibria of the static game of complete information, $\{N, \mathbf{A}, \boldsymbol{\pi}\}$ and consistent evidential equilibria of the evidential game $\Gamma = \{N, \mathbf{A}, \boldsymbol{\pi}, \mathbf{P}\}$. 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:* $\mathbf{P}_i\left(.|\sigma_i\right) = \boldsymbol{\sigma}_{-i}^*$*,* $i \in N$*. 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 $\pi_i(\sigma_i, \boldsymbol{\sigma}^*_{-i})$ 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 6 and 7) 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 8 and 9). 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 5).

(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 6 and 7). But $\mathbf{P}_i(.|\sigma_i) = \boldsymbol{\sigma}^*_{-i}$, $i \in N$, hence σ_i^* maximizes $\pi_i(\sigma_i, \boldsymbol{\sigma}^*_{-i})$ 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 to illustrate some of the key concepts developed so far. 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 ex-ante beliefs turn out to be wrong ex-post. (4) Both players use causal reasoning and ex-ante 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.

Consider the symmetric prisoners' dilemma game, where the entries in the payoff matrix, below, are payoffs; the first and the second entries are the payoffs of the row and column player respectively. Both players have two strategies C (cooperate) and D (defect).

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

Here $N = \{1, 2\}$, $A_1 = A_2 = \{C, D\}$, $\mathbf{A} = \{C, D\} \times \{C, D\}$ and $\boldsymbol{\pi}$ is given by the above payoff matrix. Each player has a dominant action, D that gives a higher payoff than C irrespective of the action of the other player. 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 on pure strategies.

Recall from Definition 2 that, in general, $P_{ij}(.|\sigma_i)$ will vary with σ_i. However, if player i uses causal reasoning then $P_{ij}(.|\sigma_i)$ will be independent of σ_i.

Case 1. $P_{12}(C|C) = 1$, $P_{12}(D|D) = 1$, $P_{21}(C|C) = 1$, $P_{21}(D|D) = 1$

Both players use evidential reasoning, so this is an evidential game. In particular, each player uses his identity social projection function (Definition 3). Together, they form an ingroup (in fact, a perfect ingroup; see Definition 4). For player 1, given his social projection function, strategy C is the unique optimal strategy (see Definition 6). Similarly, C is also the optimal strategy for player 2. Hence, (C, C) is an evidential equilibrium (Definition 7). 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 8). Hence, (C, C) is a consistent evidential equilibrium (Definition 9). In contrast, (C, C) is not the Nash equilibrium of the game. Indeed, (C, C) involves each player playing a strictly dominated strategies. However, (C, C) is Pareto optimal. Note that 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, for a review of the evidence).

Case 2. $P_{12}(C|C) = 1$, $P_{12}(D|D) = 1$, $P_{21}(C|C) = 1$, $P_{21}(C|D) = 1$

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 7) is (C, D). It is an evidential equilibrium because each player's chosen action is optimal, given his beliefs as 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 (but not generally).

Case 3. $P_{12}(D|C) = 1$, $P_{12}(C|D) = 1$, $P_{21}(D|C) = 1$, $P_{21}(C|D) = 1$

Given these social projection functions, the unique payoff maximizing strategy for each player is to play D (Definition 6). Hence, (D, D), is the unique evidential equilibrium (Definition 7). However, (D, D), is not a mutually consistent vector of strategies (Definition

8) 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 9).

Cases 4. $P_{12}(D|C) = 1$, $P_{12}(D|D) = 1$, $P_{21}(D|C) = 1$, $P_{21}(D|D) = 1$
Both players use causal reasoning, so this is a *causal game*. Given their social projection functions, playing D is the unique optimal strategy (Definition 6) for both players. Hence (D, D) is the unique evidential equilibrium (Definition 7). Furthermore, (D, D) is a mutually consistent vector of strategies (Definition 8) 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 9). 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 4 : *Game theorists may object to the use of equilibria that are not mutually consistent. Surely, they would argue, playing such games a large number of times would ensure that inconsistent equilibria are somehow weeded out. There are several responses. First, even a Nash equilibrium need not be mutually consistent (see Case 3 in the prisoner's dilemma game above). Indeed there is no notion of beliefs in the static game of complete information,* $\{N, \mathbf{A}, \boldsymbol{\pi}\}$. *Second, the essence of a static game is that it should be played only once. Even if the same generic game is played several times, it may well be with different opponents (randomized matching of subjects in experiments is designed to achieve just that outcome). Third, the evidence reviewed in section 2 shows that the outcomes are often not mutually consistent. Fourth, people make a range of important decisions which are best thought of as one-off decisions or where in practice, repetitions are extremely infrequent. For instance, the decision to choose a course in the University, a career, a house, a marriage partner, number of children, pension plan etc.*

5. Oligopoly games

In this section, we reconsider several classical models from industrial organization, in particular, the monopoly, competitive, Cournot, Bertrand and Stackelberg models. We first give the classical formulation of these models under causal reasoning and then reconsider them from the perspective of evidential reasoning.

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

$$Q = \sum_{i=1}^{n} q_i. \tag{5.1}$$

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

$$P(Q) = A - aQ, \; a > 0, \; A > 0. \tag{5.2}$$

Also for simplicity we shall take the unit production cost (or marginal cost) of firm i to be a constant, c_i, $i = 1, 2, ..., n$, where

$$0 \le c_1 \le c_2 \le ... \le c_n < A. \tag{5.3}$$

Then 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

$$\pi_i = (P - c_i)\, q_i, \; i = 1, 2, ..., n. \tag{5.4}$$

In the light of (5.1) and (5.2), the profit of firm i, (5.4), can be written in the useful form

$$\pi_i\left(q_i, \mathbf{q}_{-i}\right) = \left(A - c_i - a\sum_{j \neq i} q_j\right) q_i - a q_i^2, \; i = 1, 2, ..., n. \tag{5.5}$$

5.1. Classical oligopoly models under causal reasoning

In this subsection, all players us causal reasoning.

5.1.1. Perfect competition

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

$$P^* = c_1. \tag{5.6}$$

Hence, from (5.2), total output is

$$Q^* = \frac{A - c_1}{a}. \tag{5.7}$$

Note that outcomes (5.6)-(5.7) 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 a benevolent planner gives all rights of production (and use of technology) to a single profit maximizing monopolist. That monopolist would use the lowest cost technology, resulting in the profit function

$$\Pi(Q) = (A - c_1)Q - aQ^2. \tag{5.8}$$

Maximizing (5.8) with respect to Q gives monopoly output[17]:

$$Q^* = \frac{A - c_1}{2a}. \tag{5.9}$$

5.1.3. Cournot oligopoly

Given $\mathbf{q}_{-i}$ it easily follows from (5.5) that the profit maximizing output for firm i is

$$\widetilde{q}_i(\mathbf{q}_{-i}) = \frac{A - c_i}{2a} - \frac{1}{2}\sum_{j \neq i} q_j. \tag{5.10}$$

The Nash equilibrium, $\mathbf{q}^*$, (Definition 1, also known as the Cournot equilibrium) is the fixed point of the function $\widetilde{\mathbf{q}}(\mathbf{q})$. This can easily be found to be[18]

$$q_i^* = \frac{A + \sum_{j \neq i} c_j - nc_i}{(n+1)a}, i = 1, 2, ..., n. \tag{5.11}$$

5.1.4. A Stackelberg leader-follower model

This is a two-stage game[19] where the leaders (firms $i = m+1, m+2, ..., n$) choose their outputs first. The followers (firms $i = 1, 2, ..., m$) then choose their outputs, given the outputs of the leaders.[20] 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:

$$0 \leq c_1 = c_2 = ... = c_n = c < A. \tag{5.12}$$

We rewrite the (reaction) function (5.10) of follower i as

$$\widetilde{q}_i(\mathbf{q}_{-i}) = \frac{A - a\sum_{j=m+1}^{n} q_j - c}{2a} - \frac{1}{2}\sum_{j=1, j \neq i}^{m} q_j, \; i = 1, 2, ..., m. \tag{5.13}$$

---

[17] 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 $Q = \frac{A-c_1}{2a}$.

[18] 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, $q_i^* = \frac{A + \sum_{j \neq i} c_j - nc_i}{(n+1)a}$, $i = 1, 2, ..., n$.

[19] Unlike all the games considered above, which are single stage games.

[20] This is a generalization of the texbook case where $n = 1$ and $m = 1$.

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 (5.11):

$$q_i^* = \frac{A - a\sum_{j=m+1}^{n} q_j - c}{(m+1)\,a}, i = 1, 2, ..., m. \tag{5.14}$$

Let $Q_L = \sum_{j=m+1}^{n} q_j$ denote the total output produced by all the leader firms. Denote the total output produced by all the follower firms by $Q_F = \sum_{i=1}^{m} q_i^*$ where q_i^* is given in (5.14). Hence,

$$Q_F = \left(\frac{m}{a(1+m)}\right)(A - c - Q_L)\,. \tag{5.15}$$

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 equally, correctly anticipating the reaction of the followers as given by (5.14). Using (5.5), the joint profit maximizing level of output of the leaders can be found by choosing Q_L in order to maximize

$$\pi\left(Q_L, Q_F\right) = \left(A - c - aQ_F\right)Q_L - aQ_L^2$$

where Q_F is given by (5.15). The optimal level of joint profit maximizing output of the leaders, Q_L^*, is then shared equally among all the $n - m$ leader.

The resulting optimal levels of outputs are:

$$q_i^* = \frac{A-c}{2a\,(m+1)}, i = 1, 2, ..., m \text{ (followers)}, \tag{5.16}$$

$$q_i^* = \frac{A-c}{2a\,(n-m)}, i = m+1, m+2, ..., n \text{ (leaders)}. \tag{5.17}$$

From (5.14) we see that the followers, naturally, condition their outputs on that of the leaders. Hence, the leaders, when taking their decisions, anticipate the effect their actions will have on the followers. This is entirely consistent with causal reasoning (recall footnote 13). Note that (5.16)-(5.17) form a Nash equilibrium, in fact a subgame perfect Nash equilibrium[21], only when $n = m + 1$ (*one* leader).[22]

5.2. Oligopoly models under evidential reasoning

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 2). We also assume that each firm regards each consumer as an outgroup member.[23] This allows us

[21] Nash in the whole game and Nash in each subgame conditional on the output chosen by the leader.

[22] 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).

[23] 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).

to focus on the consequences of evidential reasoning for the producers. This also allows us to continue assuming that price is determined by an inverse demand curve. In particular, we shall continue to assume that the market demand curve is given in (5.2).[24]

We now describe an evidential equilibrium, $\mathbf{q}^*$, with the following properties. Suppose 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 $q_i - q_i^*$ then my rival, firm j, $j \neq i$, who is like minded, is also tempted to deviate by an amount $q_j - q_j^* = \lambda_{ij}(q_i - q_i^*)$, $\lambda_{ij} \in \mathbb{R}$". Thus, for instance, if $\lambda_{ij} = 1$ the firm holds an identity social projection function (Definition 3). When $\lambda_{ij} = 0$, firm i assigns no diagnostic significance to its desire to deviate from q_i^* in forming beliefs about whether firm $j \neq i$ would also wish to deviate from its output level in the evidential equilibrium, q_j^*. By not restricting λ_{ij} to be any particular real number we allow, at this stage, a wide range of possible behaviors. We will show how the models of subsection 5.1 can be obtained by choosing suitable values for λ_{ij}, $j \neq i$

We formalize this behavior by the following social projection function (Definition 2):

$$P_{ij}(q_j|q_i) = 1 \Leftrightarrow q_j = q_j^* + \lambda_{ij}(q_i - q_i^*), \; j \neq i. \tag{5.18}$$

Proposition 2 *: (a) Given the social projection functions (5.18), the unique evidential equilibrium (Definition 7), $\mathbf{q}^*$, is characterized by the following set of simultaneous linear algebraic equations*

$$\begin{bmatrix} 2+\sum_{j\neq 1}\lambda_{1j} & 1 & ... & 1 \\ 1 & 2+\sum_{j\neq 2}\lambda_{2j} & ... & 1 \\ ... & ... & ... & ... \\ 1 & 1 & ... & 2+\sum_{j\neq n}\lambda_{nj} \end{bmatrix} \begin{bmatrix} q_1^* \\ q_2^* \\ ... \\ q_n^* \end{bmatrix} = \begin{bmatrix} \frac{A-c_1}{a} \\ \frac{A-c_2}{a} \\ ... \\ \frac{A-c_n}{a} \end{bmatrix} \tag{5.19}$$

(b) Furthermore, $\mathbf{q}^*$ is a mutually consistent vector of strategies (Definition 8) 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 (5.18) such that $\mathbf{q}^*$ is a consistent evidential equilibrium.

Proof of Proposition 2: (a) Substituting from (5.18) into (5.5) gives

$$\pi_i(q_i, \mathbf{P}_i(.|q_i)) = \left\{ A - c_i - a\sum_{j\neq i}\left[q_j^* + \lambda_{ij}(q_i - q_i^*)\right]\right\} q_i - aq_i^2, \; i = 1, 2, ..., n, \tag{5.20}$$

[24] 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.

which, after simplification, gives

$$\pi_i(q_i, \mathbf{P}_i(.|q_i)) = \left(A - c_i + aq_i^* \sum_{j \neq i} \lambda_{ij} - a \sum_{j \neq i} q_j^* \right) q_i - a \left(1 + \sum_{j \neq i} \lambda_{ij} \right) q_i^2, \; i = 1, 2, ..., n. \tag{5.21}$$

Maximizing (5.21) with respect to q_i gives the optimal (pure) strategy for firm i (Definition 6), given his social projection function (5.18):

$$q_i = \frac{A - c_i + aq_i^* \sum_{j \neq i} \lambda_{ij} - a \sum_{j \neq i} q_j^*}{2a \left(1 + \sum_{j \neq i} \lambda_{ij} \right)}, \; i = 1, 2, ..., n. \tag{5.22}$$

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

$$\left(2 + \sum_{j \neq i} \lambda_{ij} \right) q_i^* + \sum_{j \neq i} q_j^* = \frac{A - c_i}{a}, \; i = 1, 2, ..., n, \tag{5.23}$$

which can be written in the matrix form

$$\begin{bmatrix} 2 + \sum_{j \neq 1} \lambda_{1j} & 1 & ... & 1 \\ 1 & 2 + \sum_{j \neq 2} \lambda_{2j} & ... & 1 \\ ... & ... & ... & ... \\ 1 & 1 & ... & 2 + \sum_{j \neq n} \lambda_{nj} \end{bmatrix} \begin{bmatrix} q_1^* \\ q_2^* \\ ... \\ q_n^* \end{bmatrix} = \begin{bmatrix} \frac{A - c_1}{a} \\ \frac{A - c_2}{a} \\ ... \\ \frac{A - c_n}{a} \end{bmatrix} \tag{5.24}$$

(b) From (5.18) we see that $P_{ij}(q_j|q_i^*) = 1 \Leftrightarrow q_j = q_j^*$. Hence, $\mathbf{q}^*$ is a mutually consistent vector of strategies and, hence, a consistent evidential equilibrium.
(c) Rewrite (5.23) in the form

$$\sum_{j \neq i} \lambda_{ij} = \frac{A - c_i}{aq_i^*} - 2 - \frac{1}{q_i^*} \sum_{j \neq i} q_j^*, \; i = 1, 2, ..., n. \tag{5.25}$$

(5.25) has many solutions, for example

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

As an application of Proposition 2 we now show how the models of subsection 5.1 can be obtained by choosing suitable values for λ_{ij}, $j \neq i$, in (5.19).

5.2.1. Perfect competition

Setting $c_1 = c_2 = ... = c_n$ and $\lambda_{ij} = -\frac{1}{n-1}$, $i \neq j$, in (5.19) gives $\sum_{i=1}^n q_i^* = \frac{A - c_1}{a}$, which equals the perfectly competitive industry output given in (5.7). Here each firm regards every other firm as an ingroup member (Definition 2c and Remark 1). The set of firms forms an ingroup (Definition 2d). 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 3.2 and, in particular, is an illustration of the contrast effect.

5.2.2. Monopoly

Setting $c_1 = c_2 = ... = c_n$ and $\lambda_{ij} = 1$, $i \neq j$, in (5.19) gives $q_i^* = \frac{A-c_1}{2na}$, $i = 1, 2, ..., n$. Thus, the aggregate output equals the level of output produced by a monopolist, and given in (5.9). The social projection functions for the producers here are identity social projection functions on the set of all producers (Definition 3). The set of firms form a perfect ingroup (Definition 4). They behave harmoniously or cooperatively towards each other (see subsection 3.2).

5.2.3. Cournot oligopoly

Setting $\lambda_{ij} = 0$, $i \neq j$, in (5.19) gives $q_i^* = \frac{A+\sum_{j\neq i} c_j - nc_i}{(n+1)a}$, $i = 1, 2, ..., n$. Thus, each firm produces the same level of output as produced by a firm under Cournot competition, as given in (5.11). Here each firm regards every other firm as an outgroup member (Definition 2c). Thus each firm regards every other player (whether consumer or producer) as an outgroup member. Hence every firm uses causal reasoning (Definition 2f and Remark 1).

5.2.4. Stackelberg leader-follower model

Let $c_1 = c_2 = ... = c_n$. Suppose we adopt the following values for λ_{ij} :

$$\lambda_{ij} = \begin{cases} 0 & if \quad i = 1, 2, ..., m,\ j = 1, 2, ..., n,\ i \neq j, \\ 1 & if \quad i = m+1, m+2, ..., n,\ j = m+1, m+2, ..., n,\ i \neq j, \\ -\frac{n-m}{m+1} & if \quad i = m+1, m+2, ..., n,\ j = 1, 2, ..., m, \end{cases} \tag{5.26}$$

From the first row of (5.26), the social projection functions for the followers (firms $1, 2, ..., m$) are all constant, hence all the followers use causal reasoning (Definition 2f and Remark 1). Each follower regards each leader as an outgroup member (Definition 2c). Hence the leaders form an outgroup relative to the followers (Definition 2e).

From the second row of (5.26), 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 3). Hence the leaders form a perfect ingroup (Definition 4).

From the third row of (5.26), each leader regards each follower as an ingroup member (Definition 2c). For instance, when $m = 1$ and $n = 2$ the leader believes that if he increased output from the evidential equilibrium, the follower will respond with contracting his output by one half of that amount. This is consistent with the follower's reaction function in (5.10). We may say that the leaders behave *collusively* towards each other but *competitively* towards the followers (the contrast effect, recall subsection 3.2).

Substituting (5.26) in (5.19) gives

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

These output levels are identical to those produced in the Stackelberg game, given in (5.16) and (5.17). However, unlike the Stackelberg game of subsection 5.1.4 (which was a two-stage game), this version is a single-stage game.

The following is a simple corollary of Proposition 2, which is of interest in its own right. It explores the effect on profits of firms as λ_{ij} varies.

Corollary 1 *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 (5.18):*
(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

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).[25] However, under evidential reasoning (subsection 5.2), $\mathbf{q}^*$ is always a consistent evidential equilibrium.

This section examines the empirical evidence (subsection 5.3), and shows that under random matching, experimental subjects' behavior does not robustly conform to a

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

Cournot-Nash equilibrium. In contrast, 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. Insofar as the experimental results appear quite clean, this range of behaviors suggest that the Cournot-Nash equilibrium is very restrictive. The outcome of oligopoly games under evidential reasoning that we have examined in section 5.2 above is able to demonstrate the range of psychological richness in behaviors.

In the early experiment 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 (an oligopoly market with two firms). The PT listed the outputs of each firm on the two margins while individual cells of the table contained the profit figures 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 profit for a given pair of quantities chosen by both firms. In recent years several experiments also provide 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 Corollary 1) in PT and PC as compared to BRO. They find that, at least once in the 20 rounds, 62%, 78% and 29% of the markets, respectively, reach the collusive outcome (or joint profit maximizing output level) 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.[26] Consider a representative paper by Huck et al. (1999), which uses symmetric firms and linear demand curves. On p. 750 they claim that when subjects are matched randomly in 10 rounds of a Cournot game, the "experimental results confirm the theory very well".[27] We find these claims to be possibly overstated. A profit table (PT) is given in which outputs of both firms vary between 3-15 (see Appendix B in Huck et al., 1999). 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. Profits are also low for an output of 3. These leaves only 7 levels of output to choose from (4,5,6,7,8,9). For round 9, under random

[26]The interested reader can consult the bibliography of Requate and Waichman (2011), which we refer to below.

[27]These views are echoed in the meta study by Huck et al. (2004).

matching, whose results the authors present as the most supportive of their hypothesis (see Table 5 in their paper), they get the following result.

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

The mean quantity is close to 8. However, the Cournot-Nash equilibrium is a prediction about the choice of individuals, not the overall average. About 65% of individuals do not choose the Cournot output level.

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 has 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 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) but the authors correctly take this as a refutation rather than a confirmation of the Cournot-Nash equilibrium.

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.

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.

6. Conclusions

In static full information games, players are uncertain about which actions the others will take. Aumann and Brandenburger (1995) gave epistemic conditions under which the play of a game would result in a Nash equilibrium (NE). A very large number of experimental subjects do not play a NE in well known games such as prisoners' dilemma, voting games, public goods games and oligopoly games. It would seem that this constitutes strong grounds for game theorists to be open to alternatives.

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 (disallowed by the Aumann-Brandenburger conditions). An example of such reasoning, called evidential reasoning (ER) in the context of the prisoners' dilemma game, where each player thinks that the other is like minded, is as follows. "I take my own preference to cooperate in a prisoner's dilemma game as being of diagnostic significance in guessing that my opponent is also likely to cooperate". So both players cooperate. 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 full 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.[28] 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 the prisoners' dilemma game is played 50-75% 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 often the collusive one, which is not a NE. 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. All these factors may be very important but ER provides yet another important argument for cooperative

[28] The formal application of our framework to other games considered in Section 3 should be obvious, although due to space limitations we have not chosen to do so.

behavior.

It is also likely that there could be 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 of an n-player Stackelberg game and the prisoner's dilemma game we allow for such a mixture. Future research, both empirical and theoretical, could fruitfully explore the idea of such mixtures.

Our framework can be extended to dynamic games and games of incomplete information; 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 thank Andrew Colman for useful comments.

References

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

[2] Aumann, R. (1976), Agreeing to Disagree. *Annals of Statistics*, 4, 1236-1239.

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

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

[5] Darai, D., and Grätz, S. (2010) Determinants of Successful Cooperation in a Face-to-Face Social Dilemma, University of Zurich, mimeo.

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

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

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

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

[10] Fouraker, L., Siegel, S. (1963). Bargaining Behavior. McGraw-Hill: New York.

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

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

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

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

[15] Huck, S., Normann, H. T., & Oechssler, J. (2004). Two are few and four are many: number effects in experimental oligopolies. *Journal of Economic Behavior and Organization*, 53, 435–446.

[16] Kay, A. C., and Ross, L. (2003). The perceptual push: The interplay of implicit cues and explicit situational construal in the Prisoners' Dilemma. *Journal of Experimental Social Psychology*, 39, 634-643.

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

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

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

[20] 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.

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

[22] Lewis, D. K., (1969). *Conventions: A Philosophical Study*. Cambridge: Harvard University Press.

[23] Loewenstein G., O'Donoghue, T., and Rabin, M. (2003). Projection Bias In Predicting Future Utility, *The Quarterly Journal of Economics*, 118: 1209-1248.

[24] Mullen, B., Atkins, J. L., Champion, D. S., Edwards, C., Hardy, D. Story, J. E., and Venderklok, M., (1985). The False Consensus Effect: A Meta–Analysis of 115 Hypothesis Tests; *Journal of Experimental Social Psychology*, 21, 263-83.

[25] Nash, J. F., (1951) Non-Cooperative Games. Annals of Mathematics 54, 286-295.

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

[27] 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.

[28] Quattrone, G. A., and Tversky, A. (1988). Contrasting rational and psychological analyses of political choice. *American Political Science Review*, 82: 719-736.

[29] 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.

[30] 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 and Organization, 41: 117-146.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.