

Isaac Levi on Rationality, Deliberation and Prediction

Cyril Hédoin

Full Professor of economics, University of Reims Champagne-Ardenne (France)

Contact: cyril.hedoin@univ-reims.fr

August 2016

This note presents and discusses Isaac Levi's account of the relationship between deliberation and prediction in decision theory. The latter is an essential part of Levi's more general theory of rationality. Levi's views potentially have tremendous implications for economists especially regarding the current use of game theory. These views are more particularly developed in several essays collected in his book *The Covenant of Reason* [1], especially "Rationality, prediction and autonomous choice", "Consequentialism and sequential choice" and "Prediction, deliberation and correlated equilibrium". The first section presents and discusses Levi's main thesis that "deliberation crowds out prediction". The next two sections will discuss some implications of this thesis for decision theory and game theory, specifically the equivalence between games in dynamic form and in normal form and the relevance of the correlated equilibrium concept for Bayesianism in the context of strategic interactions.

1. Deliberation Crowds Out Prediction

The determination of principles of rational choice is the main subject of decision theory since its early development at the beginning of the 20th century. Since its beginnings, decision theory has pursued two different and somehow conflicting goals: on the one hand, to describe and explain how people actually make choices and, on the other hand, to determine how people *should* make choices and what choices they should make. While the former goal corresponds to what can be called "positive decision theory", the latter is constitutive of "normative decision theory". Most decision theorists, especially the proponents of "Bayesian" decision theory, have agreed that decision theory cannot but be partially normative. Indeed, while today Bayesian decision theory is generally not regarded as an accurate account of how individuals are actually making choices, most decision theorists remain convinced that it is still relevant as a normative theory of rational decision-making. This is in this context that Isaac Levi's claim that "deliberation crowds out prediction" should be discussed.

In this section, I will confine the discussion to the restrictive framework of Bayesian decision theory though Levi's account more generally applies to any form of decision theory that adheres to *consequentialism*. Consequentialism will be more fully discussed in the second post of this series. Consider any decision problem \mathbf{D} in which an agent DM has to make a choice over a set of options whose consequences are not necessarily fully known for sure. Bayesians will generally model \mathbf{D} as a triple $\langle A, S, C \rangle$ where A is the set of *acts* a , S the set of *states of nature* s and C the set of *consequences* c . In the most general form of Bayesian decision theory,

any a , s and c may be regarded as a proposition to which truth-values might be assigned. In Savage's specific version of Bayesian decision theory, acts are conceived as functions from states to consequences, i.e. $a: S \rightarrow C$ or $c = a(s)$. In this framework, it is useful to see acts as DM's objects of choice, i.e. the elements over which he has a direct control, while states may be interpreted as every features in \mathbf{D} over which DM has no *direct* control. Consequences are simply the result of the combination of an act (chosen by DM) and a state (not chosen by DM). Still following Savage, it is standard to assume that DM has (subjective) beliefs over which state s actually holds. These beliefs are captured by a probability function $p(\cdot)$ with $\sum_s p(s) = 1$ for a finite state space. Moreover, each consequence c is assigned a utility value u representing DM's preferences over the consequences. A Bayesian DM will then choose the act that maximizes his expected utility given his subjective beliefs and his preferences, i.e.

$$\text{Max}_a Eu(a) = \sum_s p(s|a)u(a(s)) = \sum_s p(s|a)u(c).$$

Two things are worth noting. First, note that the probabilities that enter into the expected utility computation are *conditional* probabilities of states given acts. We should indeed account for the possibility that the probabilities of states depend on the act performed. The nature of the relationship between states and acts represented by these conditional probabilities is the main subject of conflict between *causal* and *evidential* decision theorists. Second, as it is well-known, in Savage's version of Bayesian decision theory, we start with a full ordering representing DM's preferences over *acts* and given a set of axioms, it is shown that we can derive a unique probability function $p(\cdot)$ and a cardinal utility function $u(\cdot)$ unique up to any positive affine transformation. It is indeed important to recognize that Savage's account is essentially behaviorist because it merely shows that given the fact that DM's preferences and beliefs satisfy some properties, then his choice can be *represented* as the maximization of some function with some uniqueness property. Not all Bayesian decision theorists necessarily share Savage's behaviorist commitment.

I have just stated that in Savage's account, DM ascribes probabilities to states, utilities to consequences and hence expected utilities to acts. However, if acts, states and consequences are all understood as propositions (as argued by Richard Jeffrey and Levi among others), then there is nothing in principle prohibiting to ascribe utilities to states and probabilities to both consequences *and* acts. This is this last possibility (ascribing probabilities to acts) that is the focus of Levi's claim that deliberation crowds out prediction. In particular, does it make sense for DM to have *unconditional* probabilities over the set A ? How having such probabilities could be interpreted from the perspective of DM's deliberation in \mathbf{D} ? If we take a third person perspective, ascribing probabilities to DM's objects of choice seems not particularly contentious. It makes perfect sense for me to say for instance "I believe that you will start again to smoke before the end of the month with probability p ". Ascribing probabilities to others' choices is an essential part of our daily activity consisting in predicting others' choices. Moreover, probability ascription may be a way to explain and rationalize others' behavior. The point of course is that these are *my* probabilities, not *yours*. The issue here is whether a deliberating agent has to, or even can ascribe such probabilities to his own actions, acknowledging that such probabilities are in any case not relevant in the expected utility computation.

Levi has been (with Wolfgang Spohn) the most forceful opponent to such a possibility. He basically claims that the principles of rationality that underlie any theory of decision-making (including Bayesian ones) cannot at the same time serve as explanatory and predictive tools and as normative principles guiding rational behavior. In other words, as far as the deliberating agent is using rationality principles to make the best choice, he cannot at the same time use these principles to predict his own behavior *at the very moment he is making his choice*.¹ This is the essence of the “deliberation crowds out prediction” slogan. To understand Levi’s position, it is necessary to delve into some technical details underlying the general argument. Wlodek Rabinowicz [2] makes a great job in reconstructing this argument (see also [3]). A crucial premise is that, following De Finetti, Levi considers belief ascription as fully constituted by the elicitation of betting rates, i.e. DM’s belief over some event E is determined and corresponds to what DM would consider as the fair price of a gamble where event E pays x \$ and event non- E pays y \$.² Consider this example: I propose you to pay y \$ (the cost or the price of the bet) to participate to the following bet: if Spain win the Olympic gold medal of basketball at Rio this month, I will pay you x \$, otherwise I pay you nothing. Therefore, x is the net gain of the bet and $x+y$ is called the *stake* of the bet. Now, the fair price y^* \$ of the bet corresponds to the amount for which you are indifferent between taking and not taking the bet. Suppose that $x = 100$ and that $y^* = 5$. Your betting rate for this gamble is then $y^*/(x+y^*) = 5/105 = 0,048$, i.e. you believe that Spain will win with probability less than 0,05. This is the traditional way beliefs are determined in Bayesian decision theory. Now, Levi’s argument is that such a procedure cannot be applied in the case of beliefs over acts on pain of inconsistency. The argument relies on two claims:

- (1) If DM is certain that he will not perform some action a , then a is not regarded as part of the feasible acts by DM.
- (2) If DM assigns probabilities to acts, then he must assign probability 0 to acts he regards as inadmissible, i.e. which do not maximize expected utility.

Clearly, (1) and (2) entail together that only feasible acts (figuring in the set A) are admissible (maximize expected utility), in which case deliberation is unnecessary for DM. If it is the case however, that means that principles of rationality cannot be used as normative principles in the deliberation process. While claim (1) is relatively transparent (even if it is disputable), claim (2) is less straightforward. Consider therefore the following illustration.

DM has a choice between two feasible acts a and b with $Eu(a) > Eu(b)$, i.e. only a is admissible. Suppose that DM assigns probabilities $p(a)$ and $p(b)$ according to the procedure presented above. We present DM with a fair bet B on a where the price is y^* and the stake is $x+y^*$. As the

¹ Note that this has far reaching implications for moral philosophy and ethics as *moral* decision problems are a strict subset of decision problems. All moral decision problems can be represented along a weakly consequentialist frame.

² A gamble’s fair price is the price at which DM is indifferent accepting to buy the bet and accepting to sell the bet.

bet is fair, y^* is the fair price and $y^*/(x+y^*) = p(a)$ is the betting rate measuring DM's belief. Now, DM has four feasible options:

Take the bet and choose a (B& a)

Do not take the bet and choose a (notB& a)

Take the bet and choose b (B& b)

Do not take the bet and choose b (notB& b)

As taking the bet and choosing a guarantee a sure gain of x to DM, it is easy to see that B& a strictly dominates notB& a . Similarly, as taking the bet and choosing b guarantee a sure loss of y^* , notB& b strictly dominates B& b . The choice is therefore between B& a and notB& b and clearly $Eu(a) + x > Eu(b)$. It follows that the fair price for B is $y^* = x + y^*$ (i.e. $x = 0$ for any y^*) and hence $p(a) = 1$ and $p(b) = 1 - p(a) = 0$. The inadmissible option b has probability 0 and is thus regarded as unfeasible by DM (claim 1). No deliberation is needed for DM if he predicts his choice since only a is regarded as feasible.

Levi's argument is by no means undisputable and the papers of Rabinowicz and Joyce referred above make a great job at showing its weaknesses. In the next two sections, I will however take it as granted and discuss some of its implications for decision theory and game theory.

2. Consequentialism and Sequential-Form Decision Problems

In his article "Consequentialism and sequential choice", Isaac Levi builds on his "deliberation crowds prediction" thesis to discuss Peter Hammond's account of consequentialism in decision theory presented in the paper "Consequentialist Foundations for Expected Utility" [4]. Hammonds contends that consequentialism (to be defined below) implies several properties for decision problems, especially (i) the formal equivalence between decision problems in sequential (or extensive) form and strategic (or normal) form and (ii) ordinality of preferences over options (i.e. acts and consequences). Though Levi and Hammonds are essentially concerned with one-person decision problems, the discussion is also relevant from a game-theoretic perspective as both properties are generally assumed in the latter. This section will focus on point (i).

First, what is consequentialism? Levi distinguishes between three forms: weak consequentialism (WC), strong consequentialism (SC) and Hammond's consequentialism (HC). According to Levi, while only HC entails point (i), both SC and HC entail point (ii). Levi contends however that none of them is defensible once we take into account the "deliberation crowds out prediction" thesis. We may define these various forms of consequentialism on the basis of the notation introduced in the preceding post. Recall that any decision problem D corresponds then to a triple $\langle A, S, C \rangle$ with A the set of acts (defined as functions from states to consequences), S the set of states of nature and C the set of consequences. A probability distribution over S is defined by the function $p(\cdot)$ and represents the decision-maker DM's

subjective beliefs while a cardinal utility function $u(\cdot)$ defined over C represents DM's preferences. Now the definitions of WC and SC are the following:

Weakly consequentialist representation - A representation of \mathbf{D} is *weakly consequentialist* if, for each $a \in A$, an unconditional utility value $u(c)$ is ascribed to any element c of the subset $C_a \subseteq C$ and where we allow for $a \in C_a$. If a is not the sole element of C_a , then the representation is *nontrivially* weakly consequentialist.

(WC) Any decision problem \mathbf{D} has a weakly consequentialist representation.

Strongly consequentialist representation - A representation of \mathbf{D} is *strongly consequentialist* if, (i) it is nontrivially weakly consequentialist and (ii) given the set of consequence-propositions C , if c_a and c_b are two identical propositions, then the conjuncts $a \wedge c_a$ and $b \wedge c_b$ are such that $u(a \wedge c_a) = u(b \wedge c_b)$.

(SC) Any decision problem \mathbf{D} has a strongly consequentialist representation.

WC thus holds that it is always possible to represent a decision problem as a set of acts to which we can ascribe unconditional utility value to all consequences each act leads to, and where an act itself can be analyzed as a consequence. As Levi notes, WC formulated this way is undisputable.³ SC has been endorsed by Savage and most contemporary decision theorists. The difference with WC lies in the fact that SC holds a strict separation between acts and consequences. Specifically, the utility value of any consequence c is independent of the act a that brought it. SC thus seems to exclude various forms of "procedural" account of decision problems. Actually, I am not sure that the contrast between WC and SC is as important as Levi suggests for all is required for SC is to have a sufficiently rich set of consequences C to guarantee the required independence.

According to Levi, HC is stronger than SC. This is due to the fact that while SC does not entail that sequential form and strategic form decision problems are equivalent, HC makes this equivalence its constitutive characteristic. To see this, we have to refine our definition of a decision problem to account for the specificity of the sequential form. A sequential decision problem \mathbf{SD} is constituted by a set N of nodes n with a subset $N(D)$ of decision nodes (where DM makes choice), a subset $N(C)$ of chance nodes (representing uncertainty) and a subset $N(T)$ of terminal nodes. All elements of $N(T)$ are consequence-propositions and therefore we may simply assume that $N(T) = C$. $N(D)$ is itself partitioned into information sets I where two nodes n and n' in the same I are indistinguishable for DM. For each $n \in N(D)$ DM has subjective beliefs measured by the probability function $p(\cdot | I)$ that indicates DM's belief of being at node n given that he knows I . The conditional probabilities $p(\cdot | I)$ are of course generated on the basis of the unconditional probabilities $p(\cdot)$ that DM holds at each node $n \in N(C)$. The triple $\langle N(D), N(C), N(T) \rangle$ defines a tree \mathbf{T} . Following Levi, I will however simplify the discussion by assuming perfect information and thus $N(C) = \emptyset$. Now, we define a behavior norm $B(\mathbf{T}, n)$ for

³ Note that this has far reaching implications for moral philosophy and ethics as *moral* decision problems are a strict subset of decision problems. All moral decision problems can be represented along a weakly consequentialist frame.

any tree \mathbf{T} and any decision node n in \mathbf{T} the set of admissible options (choices) from the set of available options at that node. Denote $\mathbf{T}(n)$ the subtree starting from any decision node n . A strategy (or act) specifies at least one admissible option for all decision nodes reachable, i.e. $B(\mathbf{T}(n), n)$ must be non-empty for each $n \in N(D)$. Given that $N(T) = C$, we write $C(\mathbf{T}(n))$ the subset of consequences (terminal nodes) that are reachable in the subtree $\mathbf{T}(n)$ and $B[C(\mathbf{T}(n)), n]$ the set of consequences DM would regard as admissible if all elements in $C(\mathbf{T}(n))$ were directly available at decision node n . Therefore, $B[C(\mathbf{T}(n)), n]$ is the set of admissible consequences in the strategic form equivalent of \mathbf{SD} as defined by (sub)tree $\mathbf{T}(n)$. Finally, write $\varphi(\mathbf{T}(n))$ the set of admissible consequences in the sequential form decision problem \mathbf{SD} . HC is then defined as follows:

$$(HC) \quad \varphi(\mathbf{T}(n)) = B[C(\mathbf{T}(n)), n]$$

In words, HC states the kind of representation (sequential or strategic) of a decision problem \mathbf{SD} used is irrelevant in the determination of the set of admissible consequences. Moreover, since we have assumed perfect information, it is straightforward that this is also true for admissible acts which, in sequential form decision problems, correspond to exhaustive plans of actions.

Levi argues that assuming this equivalence is too strong and cannot be an implication of consequentialism. This objection is of course grounded on the “deliberation crowds out prediction” thesis. Consider a DM faced with a decision problem \mathbf{SD} with two decision nodes. At node 1, DM has the choice between consuming drug ($a1$) or abstaining ($b1$). If he abstains, the decision problem ends but if he adduces, he then has the choice at node 2 between continuing taking drugs and becoming addict ($a2$) or stopping and avoiding addiction ($c2$). Suppose that DM’s preferences are such that $u(c2) > u(b1) > u(a2)$. DM’s available acts (or strategies) are therefore $(a1, a2)$, $(a1, c2)$, $(b1, a2)$ and $(b1, c2)$.⁴ Consider the strategic form representation of this decision problem where DM has to make a choice once for all regarding the whole decision path. Arguably, the only admissible consequence is $c2$ and therefore the only admissible act is $(a1, c2)$. Assume however that if DM were to choose $a1$, he would fall prey to temptation at node 2 and would not be able to refrain from continuing consuming drugs. In other words, at node 2, only option $a2$ would actually be available. Suppose that DM knows this at node 1.⁵ Now, a sophisticated DM will anticipate his inability to resist temptation and will choose to abstain ($b1$) at node 1. It follows that a sophisticated DM will choose $(b1, a2)$ in the extensive form of \mathbf{SD} , thus violating HC (but not SC).

What is implicit behind Levi’s claim is that, while it makes perfect sense for DM to ascribe probability (including probability 1 or 0) to his future choices at subsequent decision nodes,

⁴ Acts $(b1, a2)$ and $(b1, c2)$ are of course equivalent in terms of consequences as DM will never actually have to make a choice at node 2. Still, in some cases it is essential to determine what DM would do in counterfactual scenarios to evaluate his rationality.

⁵ Alternatively, we may suppose that DM has at node 1 a probabilistic belief over his ability to resist temptation at node 2. This can be simply implemented by adding a chance node before node 1 that determines the utility value of the augmented set of consequences and/or the available options at node 2 and by assuming that DM ignores the result of the chance move.

this cannot be the case for his choice over acts, i.e. exhaustive plans of actions in the decision path. For if it was the case, then as $(a1, c2)$ is the only admissible act, he would have to ascribe probability 0 to all acts but $(a1, c2)$ (recall Levi's claim 2 in the previous section). But then, that would also imply that only $(a1, c2)$ is feasible (Levi's claim 1) while this is actually not the case.⁶ Levi's point is thus that, *at node 1*, the choices at node 1 and at node 2 are qualitatively different: at node 1, DM has to *deliberate* over the implications of choosing the various options at his disposal *given his beliefs over what he would do at node 2*. In other words, DM's choice at node 1 requires him to deliberate on the basis of a prediction about his future behavior. In turn, at node 2, DM's choice will involve a similar kind of deliberation. The reduction of extensive form into strategic form is only possible if one conflates these two choices and thus ignores the asymmetry between deliberation and prediction.

Levi's argument is also relevant from a game-theoretic perspective as the current standard view is that a formal equivalence between strategic form and extensive form games holds. This issue is particularly significant for the study of the rationality and epistemic conditions sustaining various solution concepts. A standard assumption in game theory is that the players have knowledge (or full belief) of their strategy choices. The evaluation of the rationality of their choices both in strategic and extensive form games however requires to determine what the players believe (or would have believed) in counterfactual situations arising from different strategy choices. For instance, it is now well established that common belief in rationality does not entail the backward induction solution in perfect information games or rationalizability in strategic form games. Dealing with these issues necessitates a heavy conceptual apparatus. However, as recently argued by economist Giacomo Bonanno [5], not viewing one's strategy choices as objects of belief or knowledge allows an easier study of extensive-form games that avoids dealing with counterfactuals. Beyond the technical considerations, if one subscribes to the "deliberation crowds out prediction", this is an alternative path worth exploring.

I think that Levi's example is not fully convincing however. Arguably, one may argue that since action $c2$ is assumed to be unavailable at node 2, acts $(a1, c2)$ and $(b1, c2)$ should also be regarded as unavailable. The resulting reduced version of the strategic form decision problem would then lead to the same result than the sequential form. This is not different even if we assume that DM is uncertain regarding his ability to resist temptation (see the preceding note). Indeed, the resulting expected utilities of acts would trivially lead to the same result in the strategic and in the sequential forms. Contrary to what Levi argues, it is not clear that that would violate HC.

⁶ I think that Levi's example is not fully convincing however. Arguably, one may argue that since action $c2$ is assumed to be unavailable at node 2, acts $(a1, c2)$ and $(b1, c2)$ should also be regarded as unavailable. The resulting reduced version of the strategic form decision problem would then lead to the same result than the sequential form. This is not different even if we assume that DM is uncertain regarding his ability to resist temptation (see the preceding note). Indeed, the resulting expected utilities of acts would trivially lead to the same result in the strategic and in the sequential forms. Contrary to what Levi argues, it is not clear that that would violate HC.

3. Bayesianism, Correlated Equilibrium and the Epistemic Program

In his important article “Correlated equilibrium as an Expression of Bayesian Rationality” [6], Robert Aumann argues that Bayesian rationality in strategic interactions makes correlated equilibrium the natural solution concept for game theory. Aumann’s claim is a direct and explicit answer to a short paper by Kadane and Larkey [7] which argues that the extension of Bayesian decision theory to strategic interactions leads to a fundamental indeterminacy at the theoretical level regarding how Bayesian rational players will/should play. To these authors, the way a game will be played depends on contextual and empirical features on which the theorist has few things to say. Aumann’s aim is clearly to show that the game theorist endorsing Bayesianism is not committed to such nihilism. Aumann’s paper was one of the first contributions to what is nowadays sometimes called the “epistemic program” in game theory. The epistemic program can be characterized as an attempt to characterize various solution concepts for normal- and extensive-form games (Nash equilibrium, correlated equilibrium, rationalizability, subgame perfection, ...) in terms of sufficient epistemic conditions regarding the players’ rationality and their beliefs and knowledge over others’ choices, rationality and beliefs. While classical game theory in the tradition inspired by Nash has followed a “top-down” approach consisting in determining which strategy profiles in a game correspond to a given solution concept, the epistemic approach rather follows a “bottom-up” perspective and asks what are the conditions for a given solution concept to be implemented by the players. While Levi’s essay “Prediction, deliberation and correlated equilibrium” focuses on Aumann’s defense of the correlated equilibrium solution concept, its main points are essentially relevant for the epistemic program as a whole, as I will try to show below.

Before delving into the details of Levi’s argument, it might be useful to first provide a semi-formal definition of the correlated equilibrium solution concept. Denote $A = (A_1, \dots, A_n)$ the joint action-space corresponding to the Cartesian product of the set of pure strategies of the n players in a game. Assume that each player $i = 1, \dots, n$ has a cardinal utility function $u_i(\cdot)$ representing his preferences over the set of outcomes (i.e. strategy profiles) determined by A . Finally, denote Γ some probability space. A function $f: \Gamma \rightarrow A$ defines a correlated equilibrium if for any signal γ , $f(\gamma) = a$ is a strategy profile such that each player maximizes his expected utility conditional on the strategy he is playing:

$$\text{For all } i \text{ and all strategy } a_i' \neq a_i, Eu_i(a_i | a_i) \geq Eu_i(a_i' | a_i)$$

Correspondingly, the numbers $\text{Prob}\{f^{-1}(a)\}$ define the correlated distribution over A that is implemented in the correlated equilibrium. The set of correlated equilibria in any given game is always at least as large than the set of Nash equilibria. Indeed, a Nash equilibrium is also a correlated equilibrium while correlated equilibria correspond to the convex hull of Nash equilibria. As an illustration, consider for instance the famous hawk-dove game:

		Column	
		C	D
Row	C	5 ; 5	3 ; 7
	D	7 ; 3	2 ; 2

This game has two Nash equilibria in pure strategy (i.e. [D, C] and [C, D]) and one Nash equilibrium in mixed strategy where each player plays C with probability 1/3. There are many more correlated equilibria in this game however. One of them is trivially given for instance by the following correlated distribution:

		Column	
		C	D
Row	C	0	1/2
	D	1/2	0

In his paper, Aumann establishes the following important theorem:

Aumann's theorem – For any game, if

- (i) Each player i has a probability measure $p_i(\cdot)$ over the joint-action space A ;
- (ii) The probability measure is unique, i.e. $p_1(\cdot) = \dots = p_n(\cdot)$;
- (iii) The players are Bayesian rational (i.e. maximizes expected utility) and this is common knowledge;

then, the players implement a correlated equilibrium corresponding to a function f with the correlated distribution defined by $\text{Prob}\{f^1(a)\} = p(a)$.

The theorem thus shows that Bayesian rational players endowed with common knowledge of their rationality and a common prior belief over the joint-action space must implement a correlated equilibrium. Therefore, it seems that Kadane and Larkey were indeed too pessimistic in claiming that nothing can be said regarding what will happen in a game with Bayesian decision makers.

Levi attacks Aumann's conclusion by rejecting all of its premises. Once again, this rejection is grounded on the "deliberation crowds out prediction" thesis. Actually, Levi's makes two distinct and relatively independent critics against Aumann's assumptions. The first concerns an assumption that I have left implicit while the second targets premises (i)-(iii) together. I will consider them in turn.

Implicit in Aumann's theorem is an assumption that Levi calls "ratifiability". To understand what that means, it is useful to recall that a Bayesian decision-maker maximizes expected utility using *conditional probabilities* over states *given* acts. In other words, a Bayesian decision maker has to account for the possibility that his choice may reveal and/or influence the likelihood that a given state is the actual state. *Evidential* decision theorists like Richard Jeffrey

claim in particular that it is right to see one's choice as an evidence for the truth-value of various state-propositions even in cases where no obvious causal relationship seems to hold between one's choice and states of nature. This point is particularly significant in a game-theoretic context where while the players make choice independently (in a normal-form game), some kind of correlation over choices and beliefs may be seen as plausible. The most extreme case is provided by the prisoner's dilemma which Levi discusses at length in his essay:

		Column	
		C	D
Row	C	5 ; 5	1 ; 6
	D	6 ; 1	2 ; 2

The prisoner's dilemma has a unique Nash equilibrium: [D, D]. Clearly, given the definition given above, this strategy profile is also the sole correlated equilibrium. However, from a Bayesian perspective, Levi argues that it is perfectly fine for Row to reason along the following lines:

"Given what I know and believe about the situation and the other player, I believe almost for sure that if I play D, Column will also play D. However, I also believe that if I play C, there is a significant chance that Column will play C".

Suppose that Row's conditional probabilities are $p(\text{Column plays D} | \text{I play D}) = 1$ and $p(\text{Column plays C} | \text{I play C}) = \frac{1}{2}$. Then, Row's expected utilities are respectively $Eu(D) = 2$ and $Eu(C) = 3$. As a consequence, being Bayesian rational, Row should play C, i.e. should choose to play a dominated strategy. Is there any wrong for Row to reason this way? The definition of the correlated equilibrium solution concept excludes this kind of reasoning because, for any action a_i , the computation of expected utilities for each alternative actions a_i' should be made using the conditional probabilities $u(. | a_i)$. This corresponds indeed to the standard definition of ratifiability in decision theory as put by Jeffrey [8]: "A ratifiable decision is a decision to perform an act of maximum estimated desirability relative to the probability matrix the agent thinks he would have if he finally decided to perform that act." In the prisoner's dilemma, it is easy to see that only D is ratifiable because considering to play C with the conditional probabilities given above, Row would do better by playing D; indeed, $Eu(D | \text{Columns plays C with probability } \frac{1}{2}) > Eu(C | \text{Columns plays C with probability } \frac{1}{2})$.

As Levi recognizes, the addition of ratifiability as a criterion of rational choice leads *de facto* to exclude the possibility that the players in a game may rationally believe that some form of *causal dependence* holds between their choices. Indeed, as formally shown by Oliver Board [9], Aumann's framework tacitly builds upon an assumption of causal independence but also of *common belief* in causal independence. For some philosophers and game theorists, this is unproblematic and indeed required since it is constitutive of game-theoretic reasoning (see for instance [10]). Quite the contrary, Levi regards this exclusion as illegitimate at least on a Bayesian ground.

Levi's rejection of premises (i)-(iii) are more directly related to his "deliberation crowds out prediction thesis". Actually, we may even focus on premises (i) and (iii) as premise (ii) depends on (i). Consider the assumption that the players have a probability measure over the joint-action space first. Contrary to a standard Bayesian decision problem where the probability measure is defined over a set of states that is distinct from the set of acts, in a game-theoretic context the domain of the probability measures encompasses each player's own strategy choice. In other words, this leads the game theorist to assume that each player ascribes an unconditional probability to his own choice. I have already explained why Levi regards this assumption as unacceptable if one wants to account for the way decision makers reason and deliberate.⁷ The common prior assumption (ii) is of course even less commendable in this perspective, especially if we consider that such an assumption pushes us outside the realm of strict Bayesianism. Regarding assumption (iii), Levi's complaint is similar: common knowledge of Bayesian rationality implies that each player knows that he is rational. However, if a player knows that he is rational *before* making his choice, then he already regards as feasible only admissible acts (recall Levi's claims 1 and 2). Hence, no deliberation has to take place.

Levi's critique of premises (i)-(iii) seem to extend to the epistemic program as a whole. What is at stake here is the epistemological and methodological status of the theoretical models build by game theorists. The question is the following: what is the modeler trying to establish regarding the behavior of players in strategic interactions? There are two obvious possibilities. A first one is that, as an outside observer, the modeler is trying to make sense (i.e. to describe and to explain) of players' choices *after having observed them*. Relatedly, still as an outside observer, he may try to predict players' choices before they are made. A second possibility is to game-theoretic models as tools to account for the players' reasoning process prospectively, i.e. how players' deliberate to make choice. Levi's "deliberation crowds out prediction" thesis could grant some relevance to the first possibility may not for the second. However, he contends that Aumann's argument for correlated equilibrium cannot be only retrospective but must also be prospective.⁸ If Levi is right, the epistemic program as a whole is affected by this argument, though fortunately there is room for alternative approaches, as illustrated by Bonanno's paper mentioned in the preceding post.

References

- [1] I. Levi, *The Covenant of Reason: Rationality and the Commitments of Thought*. Cambridge University Press, 1997.

⁷ The joint-action space assumption results from a technical constraint: if we want to exclude the player's own choice from the action space, we then have to account for the fact that each player has a different action space over which he forms beliefs. In principle, this can be dealt with even though this would lead to more cumbersome formalizations.

⁸ Interestingly, in a recent paper with Jacques Dreze which makes use of the correlated equilibrium solution concept [11], Aumann indeed argues that the use of Bayesian decision theory in a game-theoretic context has such prospective relevance.

- [2] W. Rabinowicz, "Does Practical Deliberation Crowd Out Self-Prediction?," *Erkenntnis*, vol. 57, no. 1, pp. 91–122.
- [3] J. M. Joyce, "Levi on Causal Decision Theory and the Possibility of Predicting One's Own Actions," *Philos. Stud.*, vol. 110, no. 1, pp. 69–102, Jul. 2002.
- [4] P. J. Hammond, "Consequentialist foundations for expected utility," *Theory Decis.*, vol. 25, no. 1, pp. 25–78, Jul. 1988.
- [5] G. Bonanno, "A dynamic epistemic characterization of backward induction without counterfactuals," *Games Econ. Behav.*, vol. 78, pp. 31–43, Mar. 2013.
- [6] R. J. Aumann, "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, vol. 55, no. 1, pp. 1–18, Jan. 1987.
- [7] J. B. Kadane and P. D. Larkey, "Subjective Probability and the Theory of Games," *Manag. Sci.*, vol. 28, no. 2, pp. 113–120, Feb. 1982.
- [8] R. C. Jeffrey, *The Logic of Decision*. University of Chicago Press, 1990.
- [9] O. Board, "The Equivalence of Bayes and Causal Rationality in Games," *Theory Decis.*, vol. 61, no. 1, pp. 1–19, Aug. 2006.
- [10] R. Stalnaker, "Knowledge, Belief and Counterfactual Reasoning in Games," *Econ. Philos.*, vol. 12, no. 2, pp. 133–163, Oct. 1996.
- [11] R. J. Aumann and J. H. Dreze, "Rational Expectations in Games," *Am. Econ. Rev.*, vol. 98, no. 1, pp. 72–86, Mar. 2008.