Journal of Economic Theory **95**, 169–185 (2000)
doi:10.1006/jeth.2000.2670, available online at http://www.idealibrary.com on IDEAL®

Quantified Beliefs and Believed Quantities

Dov Samet

Faculty of Management, Tel Aviv University, Tel Aviv, Israel
samet@post.tau.ac.il

Received August 11, 1999; final version received February 24, 2000

Beliefs about quantities are expressed by estimations. Beliefs are quantified by ascribing to them probability numbers. It is, shown that beliefs about quantities and quantified beliefs give rise to the same model, that of a type space. We study the axiom that an agent's estimation coincides with the estimation of that estimation, showing it to be weaker than the introspection axiom, according to which an agent is certain of his own probabilistic beliefs. It implies, however, that the agent is certain that he is introspective, and it is equivalent to the axioms of averaging and conditioning, which are expressed in terms of probabilistic beliefs. *Journal of Economic Literature* Classification Number: D80. © 2000 Academic Press

1. INTRODUCTION

1.1. *Two kinds of beliefs.* There are two different ways in which numbers are involved in expressing beliefs. Jane demonstrates the first one. She says that the probability of putting a human being on Mars in her lifetime is eighty percent. The numerical value is usually understood to be a quantification of the degree of her belief that a human being will be put on Mars in her lifetime. Jane is, of course, a literary figure; flesh and blood human beings do not usually make such statements, unless they are subjects in the laboratories of social scientists. Even literary figures did not speak this way until recently. One does not find statements of such form in Shakespeare, let alone older texts, for the simple reason that before the development of probability theory in recent centuries beliefs were not quantified.

The other way in which numbers are involved in expressing beliefs is much older and simpler. When we ask John how old his neighbor is, he answers "About seventy five years old." And Mary says "There were approximately one hundred and twenty guests," when asked about yesterday's party. Kim says that it feels like the temperature in the room is 18 centigrade. In all these cases people are expressing their opinion, or belief, about a quantity, although they are not necessarily certain what it is precisely. Their belief is given as a numerical estimation of the quantity. The numerical value involved is the *content* of the belief, rather than a

0022-0531/00 $35.00

degree of the belief. Unlike probabilistic statements, sentences expressing estimations are not the result of any modern theory, and therefore it is no surprise that texts from time immemorial attest their use. The following is an example from the Bible. "So there went up thither of the people *about* three thousand men: And the men of Ai smote of them *about* thirty and six[1] men." (Josh. 7:4–5)

We examine here the relationship between these two apparently different kinds of beliefs by studying them in the framework of the same model. We show that each kind can be reduced to or explained in terms of the other. In particular, probabilistic beliefs can be viewed as an equivalent description of the observable phenomenon of estimation, analogously to the way that probabilistic beliefs and utility functions are viewed as describing preferences. We delineate, first, the considerations that lead to the model in which both kinds of beliefs are studied.

1.2. *Modeling beliefs about quantities.* Any belief about quantity has two major components. First, there is some quantity described; John's neighbor's age, or the size of Albania's population, or the distance between two cities. Then, there is the belief of the agent concerning this quantity.

Consider first the modeling of the quantity in question. Suppose John's neighbor is really 82 years old. Should we identify "John's neighbor's age" with 82? As far as John is concerned, the answer is no. He does not equate his neighbor's age with 82, as he does not know his neighbor's age. In John's mind this quantity is associated with many possible values, Thus, a quantity is more like a set of values, or a variable. Put differently, John cannot describe the state of affairs in the world by specifying all the true facts, because he does not know his neighbor's age. Instead, John can come up with a list of several descriptions of the world (which differ, at least in his neighbor's age) that he considers as possible ones.

This leads to a standard set theoretic model of state space, the elements of which, the states, or states of the world, are interpreted as possible states of affairs. The quantity "John's neighbor's age" may then have different values in different states. In other words, we can think of it as a function that attaches to each state a number, which is John's neighbor's age at the state. In general, quantities about which we have beliefs are simply functions on a state space.

What then is a belief about such a quantity? Here the answer seems to be straightforward. John estimates his neighbor's age to be 75. It seems that we can plausibly identify this estimation, or belief, with the number 75. But

[1] It is quite obvious that approximating statements are used not only in cases of lack of knowledge of a precise value. They can be used even in cases of full knowledge for purposes of rounding. This may be the case with the number three thousand in the Biblical quote. But thirty six cannot be interpreted in such a way. In this paper we ignore rounding effects.

some caution is called for. Suppose we ask Mary what *she* thinks John's estimation of his neighbor's age is. For Mary, John's estimation is a quantity in the sense that we have already used this term: she does not know for sure what John's estimation is. In other words, John's estimation itself may vary with states; it is a function on the state space. But suppose that only John's estimation is of interest to us. Since John has no doubt concerning his estimation, would it not be simpler to assume that his estimation is just a single number? Even if we answer this question in the affirmative, it is better to consider John's estimation as a quantity that depends on the state of the world, and then find the required assumptions concerning estimations that make them known to him. In summary, a quantity is modeled as a function on a state space, and a belief concerning this quantity is itself a quantity, that is, a function on the state space.

1.3. *Modeling quantified beliefs.* Looking at the difference between beliefs about quantities and quantified beliefs from a grammatical point of view, we may say that estimations are about quantities which are nouns, while quantified beliefs are about whole sentences, or statements. One believes to some degree that in ten years human beings will inhabit the moon. This belief is about the statement "In ten years human beings will inhabit the moon." In probability theory, in logic, and in economics, statements are modeled set theoretically by events, that is, subsets of a given set. Like statements, events are true or false; an event is true in all the points it contains and false in the others. The events to which a point in this set belongs correspond to the statements that hold true in this point. Thus, a point in the set specifies the state of the world, and the set itself is just the state space that we have considered in the previous subsection. Claiming that one believes to some degree in a statement could be modeled, in principle, as this degree. But if we want to make beliefs the subject of beliefs—of others' or of the same agent—then the beliefs themselves should be considered as events. Thus, that Jane believes that the probability of putting a human being on Mars in her lifetime is eighty percent, is itself an event, that is, it is true in some states and false in others. Therefore, quantified beliefs transform statements into other statements. We conclude that both quantified beliefs and beliefs about quantities can be modeled by the same basic model, that of a state space.

1.4. *High-order beliefs.* Both kinds of beliefs give rise to high-order beliefs. That is, estimation and belief operators can be applied iteratively. Thus, it is possible to estimate estimations, or more generally, quantities that are formed algebraically from estimations. Likewise, belief operators can be applied to events that are defined in terms of these operators. The iterative application of belief operators of different agents gives rise to the notion of common belief which serves as an approximation to common knowledge (Monderer and Samet [13]). Iteration of estimation operators of several agents was studied in Samet [12], in order to give a necessary and sufficient condition for the existence of a common prior.

Here we study iterations of operators of the same agent. We focus on two axioms of this type. One, which we call the *consistent estimation* axiom, says that an agent's estimation coincides with the agent's estimation of that estimation. The other, called the *introspection axiom*, says that an agent is certain of his own beliefs. Models of information in economics and game theory almost invariably make the assumption that agents are introspective. It turns out that consistent estimation does not imply introspection. It only implies that the agent is *certain* that he is introspective; it allows for the possibility that the agent is erroneously certain he is introspective. The consistent estimation axiom is also equivalent to the *averaging* and *conditioning* axioms, which are expressed in terms of belief operators. The averaging axiom says that the belief of an agent, concerning some event A, is the average of all his believed beliefs concerning A. The conditioning axiom says that the belief of the agent, concerning an event A, is the same as his conditional belief concerning A, when the condition is his belief in A itself.

In the next section we give formal definitions of state spaces, and estimation and belief operators. In Sections 3 and 4 we list and discuss non-iterative axioms for the estimation and belief operators. In Section 5 we define type functions, and show that an operator of either kind satisfies the axioms, if and only if it is derived from a type function on the state space. In Section 6 we discuss the introspection axiom. The axioms of consistent estimation, averaging and conditioning, are presented and shown to be equivalent in Section 7. In Section 8 we discuss the related literature. The proofs are in Section 9.

2. THE MODEL

Our primitive building block is a *state space*, which is a measurable space (Ω, Σ), where Σ is a σ-field on Ω. Each point ω in Ω is called a *state* and is interpreted as a full specification of the state of affairs. We call the measurable sets—the elements of Σ—*events*. Events are interpreted as set theoretic representations of statements. An event is said to be true in all the states it contains. Conversely, the events to which a state ω belongs correspond to the statements that are true in that state of the world which ω represents.

A *quantity* is a bounded measurable function on (Ω, Σ). Thus, a quantity f defines at each state ω a specific numeric value $f(\omega)$. The set of all such functions is denoted by $B(\Omega, \Sigma)$.

Definition 1. An *estimation operator* is a function $S: B(\Omega, \Sigma) \rightarrow B(\Omega, \Sigma)$.

For given quantity f and state ω, $Sf(\omega)$ is the agent's estimation of the quantity f at ω.

Definition 2. A family of *belief operators* is a family of functions

$$B^p: \Sigma \rightarrow \Sigma,$$

for all real numbers $p \geqslant 0$.

For a given event A, $B^p(A)$ is itself an event which we interpret as the statement that the degree of the agent's belief in A is at least p.

Estimation and quantified belief operators are as yet arbitrary functions. In the next two sections we consider desired properties of, or axioms on, these operators.

3. BELIEFS ABOUT QUANTITIES

Consider the following four assumptions on the way people estimate quantities. The assumptions are quite reasonable, although we do not claim that they are universally true, or that they fully capture the complex logical and psychological mechanism of estimation.

- When the value of a quantity is certain, then it is estimated to have this value.
- If a quantity is certain to be non-negative, then its estimation is also certain to be so. (e.g., distances are always estimated by non-negative numbers).
- The estimation of the sum of two quantities is the sum of the estimations of each of them (e.g., the estimated number of guests in yesterday's party is the sum of the estimations of the number of men and women).
- The smaller a quantity is certain to be, the smaller its estimation is certain to be.

The notion certainty in this list of properties is used in its everyday sense and not in the probabilistic sense. Agents are certain of a fact if they think it is true. In terms of our model, a fact which is true in all states is certain. Thus, when a quantity has a fixed value c in all states, then it is certain to have this value. The first property says that in such a case, the agent's estimation would be c in all the states. The next two properties are likewise translated into the formal model. We interpret the last axiom as expressing continuity of the estimation operator, where the topology on $B(\Omega, \Sigma)$ is defined by uniform convergence.

We thus introduce the following four axioms on estimation operators.

$$Sc = c \tag{S1}$$

$$f \geqslant 0 \Rightarrow Sf \geqslant 0 \tag{S2}$$

$$S(f+g) = Sf + Sg \tag{S3}$$

$$f_n \to 0 \Rightarrow Sf_n \to 0. \tag{S4}$$

4. PROBABILISTIC BELIEFS

We list below seven axioms on the family of belief operators B^p that render the beliefs probabilistic. For an event A, we denote by $\neg A$ the complement of A. Thinking of events as representing statements, the set theoretic operations $\neg$, $\cap$, and $\cup$ represent the logical connectors "not", "and" and "or", respectively.

$$B^0(A) = \Omega \tag{B1}$$

$$B^1(\Omega) = \Omega \tag{B2}$$

$$p > 1 \Rightarrow B^p(A) = \varnothing \tag{B3}$$

$$p_n \uparrow p \Rightarrow B^{p_n}(A) \downarrow B^p(A) \tag{B4}$$

$$B^p(A \cap F) \cap B^q(A \cap \neg F) \subseteq B^{p+q}(A) \tag{B5}$$

$$\neg B^p(A \cap F) \cap \neg B^q(A \cap \neg F) \subseteq \neg B^{p+q}(A) \tag{B6}$$

$$A_n \downarrow A \Rightarrow B^p(A_n) \downarrow B^p(A) \tag{B7}$$

By the first axiom the degree of belief is always non-negative. In particular, this axiom ensures that the agent always holds some belief. The second axiom says that the degree of belief in the sure event is at least one, while the third axiom guarantees that the degree of belief never exceeds one. Axiom (B4) guarantees that degrees of belief enjoy the supremum property of real numbers. That is, if the degree of belief in some event is greater than or equal to each number in a given set of numbers, then it is also greater than or equal to the supremum of the numbers in this set. In particular, applying (B4) to the sequence $p, q, q, \ldots$, for $p < q$, implies that B^p is monotonic in p. By (B5), the degree of belief is superadditive. Note that $\neg B^p(A)$ is the event that the degree of belief in A is less than p.

Thus, (B6) states that the degree of belief is subadditive. Finally, (B7) states that beliefs are continuous, and implies in particular, that $B^p(A)$ is monotonic in A.

5. TYPE FUNCTIONS

5.1. *Defining type functions.* In this section we present a third description of an agent's beliefs in a state space. We show that it is equivalent to either of the descriptions in terms of the operators previously introduced, provided that the operators satisfy the corresponding list of axioms.

If a state is interpreted as a full description of the state of the world, then in particular it specifies the agent's beliefs, which are part of this state of affairs. A *type function* is such a specification. It associates with each state a full description of the agent's beliefs at the state, in the form of a probability measure over the state space. The formal definition follows.

DEFINITION 3. A *type function* on a state space (Ω, Σ) is a real valued function $T: \Omega \times \Sigma \to R$, which satisfies:

$$T(\omega, \cdot) \text{ is a probability measure on } (\Omega, \Sigma) \text{ for each state } \omega. \tag{1}$$

$$T(\cdot, A) \text{ is a measurable function for each event } A, \tag{2}$$

Thus, $T(\omega, A)$ is the probability of, or the degree of belief in A, at ω. The measure $T(\omega, \cdot)$ is called the *type* of the agent at ω; it fully specifies the agent's beliefs at ω. A state space endowed with a type function is called a *type space*.

5.2. *Types and estimation operators.* Using a type function T, we define for each function f in $B(\Omega, \Sigma)$ a function Tf on Ω by

$$Tf(\omega) = \int T(\omega, d\xi)\, f(\xi). \tag{3}$$

That is, the value of Tf at ω is the expected value of f with respect to the agent's type at ω.[2] The following theorem shows that type functions and estimation operators that satisfy axioms (S1)–(S4) are two equivalent ways to describe believed quantities.

[2] The expected value of f with respect to some probability measure μ is $\int f(\xi)\, \mu(d\xi)$. The probability measure in the righthand side of (3) is $T(\omega, \cdot)$, and therefore $\mu(d\xi)$ is $T(\omega, d\xi)$. The notation in (3) is standard in the literature of Markov processes, and is analogous to the product of a Markov transition matrix with a vector in the finite case (see Subsection 5.4).

THEOREM 1. *There is a one-to-one correspondence,* $T \leftrightarrow S$, *between type functions* T *on* (Ω, Σ), *and those estimation operators* S *on* $B(\Omega, \Sigma)$ *that satisfy axioms* (S1)–(S4), *such that* $Tf = Sf$ *for each* f.

5.3. *Types and belief operators.* The type function T can also be used to define belief operators on Σ by

$$\begin{aligned} B^p(A) &= T^{-1}(\cdot, A)([p, 1]) \\ &= \{\omega \mid T(\omega, A) \geqslant p\}. \end{aligned} \tag{4}$$

By (2), $B^p(A)$ is an event.

The next theorem shows that type functions and belief operators that satisfy axioms (B1)–(B7) are two equivalent ways to describe quantified beliefs.

THEOREM 2. *A family of belief operators* B^p *on* Σ *satisfies axioms* (B1)–(B7), *if and only if there exists a type function* T *on* (Ω, Σ) *such that for each* p, B^p *is given by* (4).

5.4. *Type functions as Markovian kernels.*[3] The mathematical object which we defined as a type function has been well studied in probability theory under the name *Markovian kernel.* It describes the transition probability of a Markov process, where $T(\omega, A)$ is the probability that the state will be in event A in the next period when it is now in state ω.

We have seen that T can be associated with a linear operator on $B(\Omega, \Sigma)$. It also can be associated with a linear operator on $M(\Omega, \Sigma)$, the space of all bounded measures on (Ω, Σ). For each bounded measure μ define a bounded measure μT by

$$\mu T(A) = \int \mu(d\xi)\, T(\xi, A). \tag{5}$$

A measure μ is *invariant* for T, if $\mu T = \mu$.

Using (5) we can define the product of two Markovian kernels, S and T, to be the Markovian kernel given by

$$(ST)(\omega, \cdot) = S(\omega, \cdot)\, T \tag{6}$$

or equivalently by (5),

$$(ST)(\omega, A) = \int S(\omega, d\xi)\, T(\xi, A). \tag{7}$$

[3] This subsection is required for Theorem 4 and for the proof of Theorem 5.

It can be easily shown that for each $f \in B(\Omega, \Sigma)$, $(ST)f = S(Tf)$, and thus we can write STf with no risk of confusion.

An event A is an *invariant set* of T, if $T1_A = 1_A$, or equivalently, if for each $\omega \in A$, $T(\omega, A) = 1$, and for each $\omega \in \neg A$, $T(\omega, A) = 0$. The set of all invariant sets is obviously a σ-field.

For a finite state space, a Markovian kernel can be described by a stochastic (Markovian) matrix the rows and columns of which are designated by states, and the row associated with a state ω is the probability vector $T(\omega, \cdot)$. The operators defined by the Markovian kernels on functions and measures are given by the product of the matrix with vectors from the right and from the left correspondingly.

6. INTROSPECTION

If $\omega \in B^p(A) \cap \neg B^1(B^p(A))$, then our agent's degree of belief in A, at ω, is at least p, but the agent is not certain at this state that this is indeed the case. In this state the agent, probing his own beliefs, is unable to tell them for sure. We say that the agent is *introspective* at ω when this does not happen, that is, when ω belongs to the complement of this event.

DEFINITION 4. The agent is *introspective* in state ω if for each p and A,

$$\omega \in \neg B^p(A) \cup B^1(B^p(A)).$$

The introspection axiom requires that the agent is always (i.e., in all states) introspective.

The introspection axiom.

$$\neg B^p(A) \cup B^1(B^p(A)) = \Omega. \tag{B8}$$

Equivalently, this axiom can be stated as

$$B^p(A) \subseteq B^1(B^p(A)). \tag{B8$'$}$$

Introspection has a simple form when the state space is separable, that is, when there exists a countable set of events that generate Σ. This form is the one most frequently used in economics and game theory and is known as the *partition model.* Let us denote $[\omega] = \{\omega' \mid T(\omega', \cdot) = T(\omega, \cdot)\}$.

THEOREM 3. *Let T be a type function on a separable state space (Ω, Σ). Then, for each state ω, $[\omega]$ is a measurable set. Moreover, if B^p are the* *belief operators defined by T, then the introspection axiom holds for these belief operators, if and only if for each ω,*

$$T(\omega, [\omega]) = 1. \tag{8}$$

Thus, under the separability assumption, the case that the agent is of a given type $T(\omega, \cdot)$ is a well defined event, $[\omega]$. The partition of the space according to the agent's type is measurable. Introspection is equivalent, in this case, to the requirement that at each state the agent is certain of his type.

The next theorem characterizes introspection in terms of the Markovian properties of T.

THEOREM 4. *Suppose that the belief operators are defined by the type function T. Then, the introspection axiom holds, if and only if the events $B^p(A)$ generate the σ-field of the invariant sets of T.*

7. BELIEFS ABOUT BELIEFS

7.1. *Three axioms.* In this section we discuss three axioms that, like the introspection axiom, involve repeated application of operators. The axioms, which are weaker than the introspection axiom, are shown to be equivalent, and to imply the property that the agent is certain that he is introspective.

Suppose we ask John to estimate his estimation of his neighbor's age. If John's beliefs are transparent to him, if he is certain of his estimations, then we would expect that his estimation of his estimation would be the same as his estimation. This requirement is formulated in the following axiom.

The consistent estimation axiom.

$$S(Sf) = Sf. \tag{S5}$$

Observe that (S5) does not imply that John is certain of his estimation. It only says that summing up, in one number, his uncertainty concerning his own estimation Sf, John comes up with an estimation that coincides with the estimation of f.

The next two axioms are expressed in terms of belief operators B^p, but we first discuss them in terms of operators which we denote B_0^p. For an event A, $B_0^p(A)$ is the event that the degree of belief in A is *exactly p*. When the operators B^p satisfy axiom (B1)–(B7), then $B_0^p(A)$ can be expressed in terms of these operators: $B_0^p(A) = B^p(A) \cap B^{1-p}(\neg A)$.

Suppose that in state ω each of the events $B_0^{p_i}(B_0^{r_i}(A))$ is true, where the events $B_0^{r_i}(A)$ are disjoint in pairs, and $\sum p_i = 1$. Thus the agent is uncertain, at ω, of his degree of belief concerning A. Our first axiom requires that in this case, his degree of beliefs in A is the average of the r_i's, namely, it is $r = \sum p_i r_i$. Thus the axiom says that

$$\bigcap_i B_0^{p_i}(B_0^{r_i}(A)) \subseteq B_0^r(A). \tag{9}$$

However, this axiom has very little bite, as in general there may be, for a given A, a continuum of non-empty disjoint events $B_0^r(A)$, and for any $p_i > 0$ and r_i, $B_0^{p_i}(B_0^{r_i}(A)) = \varnothing$. To overcome this difficulty we need only replace the events that describe precise degree of belief by events that describe degree of belief in an interval. For this purpose we denote by $B^{p,q}(A)$, for $p < q$, the event that the degree of belief in A is in the interval $[p, q)$, i.e., $B^{p,q}(A) = B^p(A) \cap \neg B^q(A)$.

The averaging axiom. For numbers $r_0 < r_1 < \cdots < r_n$, such that $r_0 = 0$ and $r_n > 1$, pairs $\underline{p}_i < \bar{p}_i$, for $i = 1, ..., n$, and $\underline{r} = \sum_{i=1}^n \underline{p}_i r_{i-1}$, and $\bar{r} = \sum_{i=1}^n \bar{p}_i r_i$:

$$\bigcap_{i=1}^n B^{\underline{p}_i, \bar{p}_i}(B^{r_{i-1}, r_i}(A)) \subseteq B^{\underline{r}, \bar{r}}(A), \tag{B9}$$

Thus, the events $B^{r_{i-1}, r_i}(A)$, which are pairwise disjoint, correspond to the events $B_0^{r_i}(A)$ in (9), $B^{\underline{p}_i, \bar{p}_i}$ correspond to $B_0^{p_i}$, and $\underline{r}$ and $\bar{r}$, which are the bounds for the degree of belief in A, correspond to r.

Next consider the axiom,

$$B_0^p(B_0^r(A)) \subseteq B_0^{pr}(A \cap B_0^r(A)). \tag{10}$$

It can be read as follows. If the degree of belief in $B_0^r(A)$ is p, then the degree of belief in $A \cap B_0^r(A)$ divided by p is r. This simply means that the degree of the conditional belief in A is r, when the condition is the event that the degree of the belief in A is r. Here again, the axiom cannot be stated in this form for the reason given before. To state it correctly we substitute intervals for the precise degrees of beliefs.

The conditioning axiom. For $r_1 < r_2$ and $\underline{p} < \bar{p}$:

$$B^{\underline{p}, \bar{p}}(B^{r_1, r_2}(A)) \subseteq B^{\underline{p} r_1, \bar{p} r_2}(A \cap B^{r_1, r_2}(A)). \tag{B10}$$

7.2. *The equivalence theorem.*

THEOREM 5. *Let S be an estimation operator and $\{B^p\}$ a family of belief operators, which are defined by the same type function on (Ω, Σ). Then, the* *consistent estimation, averaging, and conditioning axioms are equivalent, and they imply*

$$B^1(\neg B^p(A) \cup B^1(B^p(A))) = \Omega. \tag{11}$$

Thus, although the axioms in Theorem 5 do not imply introspection, they do imply by (11) and by Definition (4), that the agent is always certain that he is introspective.

The proof that conditioning implies averaging, and that the latter implies consistent estimation require elementary probability theory. The converse implications, as well as the implication of (11), make use of the ergodic theory of Markov kernels.

Property (11) is weaker than the three equivalent conditions in Theorem 5, as the following example demonstrates. Let $\Omega = \{1, 2, 3\}$ and $T(i, \{j\}) = 1/2$ for i and j in $\{1, 2\}$, and $T(3, \{1\}) = 1$. While (11) holds, $T^2 \neq T$.

The following theorem describes the simple structure of separable belief spaces that satisfy (11).

THEOREM 6. *Let (Ω, Σ) be a separable state space with a type function T, and let B^p be the belief operators defined by T. Then, (11) is satisfied, if and only if all the states in which the agent is introspective form a nonempty event I, and in each state the agent is certain that he is introspective, i.e., $B^1(I) = \Omega$.*

Note that although the event that the agent is introspective, I, is not empty it can be very small. Indeed, if $\{\omega_0\}$ is measurable, then, $T(\omega, \{\omega_0\}) = 1$, for each ω, defines a type function for which $I = \{\omega_0\}$.

8. RELATED LITERATURE

Harsanyi [4] introduced the idea of type spaces, in order to model the interaction of beliefs in games with incomplete information. A simpler description of type spaces, and a construction of a universal type space, were given by Mertens and Zamir [11] for topological type spaces. Heifetz and Samet [8] constructed a universal type space for general measure theoretic type spaces. In all these models introspection of agents was assumed.

The relation between type spaces and belief operators has been studied by several authors. Our characterization of type spaces in terms of these operators is similar to the one in Gaifman [3]. Fagin *et al.* [2], Aumann [1], and Heifetz and Mongin [7] studied belief operators in a framework of formal logic.

Computing the expected value of functions over type spaces, that is, the use of type functions as estimation operators, is well known. In fact, it is the main reason for using these spaces in economics and game theory.

Various forms of the consistent estimation, averaging and conditioning axioms, which are shown to be equivalent in Theorem 5, have been discussed in the literature. Several axioms have been proposed, in the spirit of the conditioning axiom. They require that the conditional probability of an event A, given event B which describes some probabilistic statements about A, is compatible with the description in B (for a short survey and references, see Halpern [5]). The various versions differ in the kind of the conditioning event, as well as in whose beliefs are described in B, and whose in the conditional belief. Halpern [5] studied, for the finite case, an axiom similar to our conditioning axiom. Correcting a claim made in this paper, Halpern [6] showed that finite spaces in which this axiom is satisfied have a structure analogous to that in Theorem 6.

Several authors consider the axioms of averaging and conditioning as a significant departure from introspection (see, e.g., Skyrms [14] and Jeffrey [9]). Skyrms [14] also showed that the conditioning axiom implies the averaging axiom, although not in a model of type spaces. As shown here, the escape from introspection is not that dramatic, since each of these equivalent axioms implies (11), which expresses certainty of introspection. Gaifman [3] studied axioms concerning the relation between a type function and a given measure μ, over the state space. In particular, he considered the condition $T^2 = T$ where the equality holds μ-almost everywhere.

The ergodic nature of type functions was first recognized by Gaifman [3], but he omitted that part of the proof relating to ergodic theory. In [12], Samet used results for discrete Markov chains to give a necessary and sufficient condition for the existence of a common prior.

9. PROOFS

Proof of Theorem 1. It is straightforward to see that T satisfies axioms (S1)–(S4). We need to show that Tf is in $B(\Omega, \Sigma)$. Obviously it is bounded. For any characteristic function 1_A, $T1_A(\omega) = T(\omega, A)$, and thus by (2), Tf is measurable for characteristic functions. By linearity of the integral Tf is measurable for step functions. Finally, for a given f let f_n be a sequence of step functions that converges to f. Then, by (S4), Tf_n converges to Tf and is therefore measurable.

For a given estimation operator S define $T(\omega, A) = S1_A(\omega)$. The measurability of $S1_A$ implies (2), and axioms (S1)–(S3) imply (1). For any fixed ω, $Sf(\omega)$ and $Tf(\omega)$ are linear functionals on $B(\Omega, \Sigma)$ which agree on characteristic functions, and therefore they coincide. Hence, $Sf = Tf$. ▮

Proof of Theorem 2. It is easy to verify that belief operators that are defined by (4) satisfy all the axioms.

Conversely, suppose that the operators B^p satisfy (B1)–(B7). Consider the set $P(\omega, A) = \{p \mid \omega \in B^p(A)\}$. By (B1) it is non-empty, and by (B4) it is a closed subset of $[0, 1]$. Moreover, as (B4) implies monotonicity, this closed set is of the form $[0, p]$. Define $T(\omega, A) = \max P(\omega, A)$. The additivity and continuity of $T(\omega, \cdot)$ follow from (B5)–(B7). By (B3) it is a probability measure. The equality in (4) follows easily from the fact that $P(\omega, A)$ is an interval. The measurability condition (2) follows from (4). ▮

Proof of Theorem 3. Let Σ_0 be a countable set that generates Σ. Then,

$$[\omega] = \bigcap_{p,\, A:\, \omega \in B^p(A)} B^p(A),$$

where the intersection is over A in Σ_0 and rational p, since $T(\omega', \cdot)$ is the same as $T(\omega, \cdot)$, iff they coincide on Σ_0. Thus, $[\omega]$ is an event. When the agent is introspective at ω, then $T(\omega, B^p(A)) = 1$ whenever $\omega \in B^p(A)$, and thus (8) holds. Conversely, since $[\omega] \subseteq B^p(A)$ for any A and p such that $\omega \in B^p(A)$, (8) implies $T(\omega, B^p(A)) = 1$, i.e., $\omega \in B^1(B^p(A))$. ▮

Proof of Theorem 4. If $B^p(A)$ is invariant then for each $\omega \in B^p(A)$, $T(\omega, B^p(A)) = 1$ and therefore (B8′) holds. Conversely, if (B8′) holds, then $T(\omega, B^p(A)) = 1$ for all $\omega \in B^p(A)$. Also, if $p_n < p$ and p_n converges to p, then

$$\begin{aligned}\neg B^p(A) &= \bigcup_n B^{1-p_n}(\neg A) \subseteq \bigcup_n B^1(B^{1-p_n}(\neg A)) \\ &= B^1\left(\bigcup_n B^{1-p_n}(\neg A)\right) \\ &= B^1(\neg B^p(A)).\end{aligned}$$

Hence, for each $\omega \in \neg B^p(A)$, $T(\omega, \neg B^p(A)) = 1$, which proves that $B^p(A)$ is invariant. If A is an invariant set, then $T(\omega, A) = 1$ for each $\omega \in A$, and $T(\omega, A) = 0$ for each $\omega \in \neg A$. Thus, $A = B^1(A)$, and A belongs to the said σ-field. ▮

Proof of Theorem 5. We need for the proof some preliminaries from the ergodic theory of Markov kernels. Let μ be a finite measure on (Ω, Σ). We say that T is *null preserving* with respect to μ, if $\mu(A)=0$ implies that $T(\omega, A)=0$, μ-almost everywhere. The ergodic theory studies the Kernel T as an operator on the space of measures that are absolutely continuous with respect to μ (a space that can be identified with $L_1(\mu)$, by the identification $v \to dv/d\mu$) and on its dual space $L_\infty(\mu)$. Note, however, that the results in Theorem 5 hold at *all* the points in Ω and not almost everywhere with respect to some measure.

Given a measure μ, the state space Ω can be decomposed, in a unique way, up to μ-null sets, into two measurable parts: the *conservative* part C, and the *dissipative* part $D=\neg C$. In the case of a discrete space these parts consist of the recurrent and transient states respectively. For our purposes we need only the following properties of the two parts.

PROPOSITION 1. 1. *There exists* $f \geqslant 0$ *in* $L_\infty(\mu)$ *such that* $\{\omega \mid f(\omega)>0\}=D$, *and* $\sum_{k=0}^{n} T^k f<1$ *for all* n.

2. $T(\omega, C)=1$, *for* μ*-almost all* $\omega \in C$.

The two parts of this proposition are Theorem 1.5 and Corollary 1.4 in Krengel ([10], Chapter 3).

An event B is *μ-invariant* if $T1_B=1_B$, as functions in $L_\infty(\mu)$, or equivalently, if $T(\omega, B)=1$ μ-almost surely on B and $T(\omega, \neg B)=0$ μ-almost surely on $\neg B$. It is easy to see that the family of all invariant sets is a σ-field. The following proposition is from Krengel ([10], Chapter 3, Lemma 3.3).

PROPOSITION 2. *Suppose that* $\Omega=C$. *Then, for* $f \in L_\infty(\mu)$, $Tf=f$, *if and only if* f *is measurable with respect to the* σ*-field of* μ*-invariant sets.*

We show first that (S5) implies (11). As $Tf=Sf$ for each f, it follows that $T^2 f=S^2 f$, and by (S5), $T^2=T$.

Fix a state ω_0 and let $\mu=T(\omega_0, \cdot)$. By the definition of the product of kernels, (6),

$$\mu T=T^2(\omega_0, \cdot)=T(\omega_0, \cdot)=\mu, \tag{12}$$

and therefore, μ is an invariant measure of T. Thus, $\mu(A)=\int \mu(d\omega)\, T(\omega, A)$. In particular, if $\mu(A)=0$, then $T(\omega, A)=0$ μ-almost surely, that is, T is null preserving with respect to μ, which justifies the use of Propositions 1 and 2.

From $T^2=T$ we conclude that for the function f in Proposition 1, $f+nTf<1$ for each n, which shows that for μ-almost all $\omega \in D$, $Tf(\omega)=0$, or equivalently, $T(\omega, C)=1$. With the second part of this proposition, we conclude that $T(\omega, C)=1$ for μ-almost all ω. Therefore $\mu(C)=\int \mu(d\omega)\, T(\omega, C)=1$, and $\Omega=C$ up to μ-null sets.

Suppose that A is a μ-invariant set. Then for μ-almost all ω in A, $\omega \in B^1(A)$. Thus, $\mu(\neg A \cup B^1(A)) = 1$. Remembering that $\mu = T(\omega_0, \cdot)$, we see that $\omega_0 \in B^1(\neg A \cup B^1(A))$. Since ω_0 was arbitrary, we conclude that

$$B^1 \neg A \cup B^1(A) = \Omega. \tag{13}$$

To prove (11) it is enough, by (13), to show that for any p and A, $B^p(A)$ is μ-invariant. This follows from Proposition 2, for $f = T1_A$, as $B^p(A) = \{\omega \mid f(\omega) \geqslant p\}$ and $Tf = T^2 1_A = f$.

We show, now, that (S5) implies (B10). Suppose $\omega_0 \in B^{\underline{p}, \bar{p}}(B^{r_1, r_2}(A))$ and let $\mu = T(\omega_0, \cdot)$. Then, $\mu(B^{r_1, r_2}(A)) \in [\underline{p}, \bar{p})$. As we have shown, μ is invariant, and therefore $\mu(A \cap B^{r_1, r_2}(A)) = \int \mu(d\omega)\, T(\omega, A \cap B^{r_1, r_2}(A))$. We have also proved that $B^{r_1, r_2}(A)$ is a μ-invariant set, and thus for μ-almost all ω in $B^{r_1, r_2}(A)$, $T(\omega, B^{r_1, r_2}(A)) = 1$ and for μ-almost all ω in $\neg B^{r_1, r_2}(A)$, $T(\omega, B^{r_1, r_2}(A)) = 0$. Thus, the last integral is $\int_{B^{r_1, r_2}(A)} \mu(d\omega)\, T(\omega, A)$. But the integrand takes values in $[r_1, r_2)$ over $B^{r_1, r_2}(A)$, and hence the integral is in $[\underline{p}r_1, \bar{p}r_2)$. Thus, $\mu(A \cap B^{r_1, r_2}(A)) \in [\underline{p}r_1, \bar{p}r_2)$, or equivalently, $\omega_0 \in B^{\underline{p}r_1, \bar{p}r_2}(A \cap B^{r_1, r_2}(A))$, which shows that (B10) holds.

To see that (B10) implies (B9) note that by (B10)

$$\bigcap_{i=1}^{n} B^{\underline{p}_i, \bar{p}_i}(B^{r_{i-1}, r_i}(A)) \subseteq \bigcap_{i=1}^{n} B^{\underline{p}_i r_{i-1}, \bar{p}_i r_i}(A \cap B^{r_{i-1}, r_i}(A)).$$

Since the sets $B^{r_{i-1}, r_i}(A)$ form a partition of the space, it follows that

$$\bigcap_{i=1}^{n} B^{\underline{p}_i r_{i-1}}(A \cap B^{r_{i-1}, r_i}(A)) \subseteq B^{\sum \underline{p}_i r_{i-1}}(A),$$

and similarly

$$\bigcap_{i=1}^{n} \neg B^{\bar{p}_i r_i}(A \cap B^{r_{i-1}, r_i}(A)) \subseteq \neg B^{\sum \bar{p}_i r_i}(A),$$

which proves (B9).

Finally, to see that (B9) implies (S5), let $\mu = T(\omega_0, \cdot)$. Choose a sequence $r_0 < r_1 < \cdots < r_n$, such that $r_0 = 0$, $r_n > 1$ and $r_i - r_{i-1} < \varepsilon$. Let $\mu(B^{r_{i-1}, r_i}(A)) \in [\underline{p}_i, \bar{p}_i)$, where $\bar{p}_i - \underline{p}_i < \varepsilon / \sum r_i$. Then, by (B9), $\mu(A)$ is in $[\underline{r}, \bar{r})$. Also,

$$\int \mu(d\omega)\, T(\omega, A) = \sum \int_{B^{r_{i-1}, r_i}(A)} \mu(d\omega)\, T(\omega, A).$$

The i term in this sum is in $[\underline{p}_i r_{i-1}, \bar{p}_i r_i)$, and therefore the integral is in $[\underline{r}, \bar{r})$. As $\bar{r} - \underline{r} < 2\varepsilon$, it follows that $\mu(A) = \int \mu(d\omega)\, T(\omega, A)$. By (7), this means that $T(\omega_0, A) = T^2(\omega_0, A)$, and therefore, $T^2 = T$. ▮

Proof of Theorem 6. Suppose that (11) holds, and let $I = \bigcap \neg B^p(A) \cup B^1(B^p(A))$, where the intersection, here and in the rest of the proof, is over events A in a countable generating family of events Σ_0, and rational p. Then, by (11), $B^1(I) = \Omega$, and in particular, I is not empty. Moreover, $\omega \in I$, iff

$$\omega \in \bigcap_{p,\, A:\, \omega \in B^p(A)} B^1(B^p(A)).$$

This holds, iff

$$\omega \in B^1\left(\bigcap_{p,\, A:\, \omega \in B^p(A)} B^p(A)\right),$$

that is, iff $T(\omega, [\omega]) = 1$. We conclude, by Theorem 3, that the agent is introspective at ω iff $\omega \in I$. The converse is obvious. ▮

REFERENCES

1. R. J. Aumann, Interactive epistemology II: probability, *Int. J. Game Theory* **28** (1999), 301–314.
2. R. Fagin, J. Y. Halpern, and N. Megiddo, A logic of reasoning about probabilities, *Info. Comput.* **87** (1990), 78–128.
3. H. Gaifman, A theory of higher order probabilities, *in* "Causation, Chance and Credence" (B. Skyrms and W. Harper, Eds.), Kluwer Academic, Dordrecht/Norwell, 1988.
4. J. C. Harsanyi, Games with incomplete information played by Bayesian players, parts I, II, & III, *Man. Sc.* **14** (1967-68), 159–182, 320–334, 486–502.
5. J. Y. Halpern, The relationship between knowledge, belief and certainty, *Ann. Math. Artificial Intelligence* **4** (1991), 301–322.
6. J. Y. Halpern, Errata: The relationship between knowledge, belief and certainty, *Ann. Math. Artificial Intelligence* **26** (1999), 253–256.
7. A. Heifetz and P. Mongin, Probability logic for type spaces, *Games Econ. Behavior*, forthcoming.
8. A. Heifetz and D. Samet, Topology-free typology of beliefs, *J. Econ. Theory*, forthcoming.
9. R. Jeffrey, "Probability and the Art of Judgment," Cambridge University Press, Cambridge, UK, 1992.
10. U. Krengel, "Ergodic Theorems," Studies in Math., Vol. 6, de Gruyter, Berlin, 1985.
11. J. F. Mertens and S. Zamir, Formulation of Bayesian analysis for games with incomplete information, *Int. J. Game Theory* **14** (1985), 1–29.
12. D. Samet, Iterative expectations and common priors, *Games Econ. Behavior* **24** (1997), 131–141.
13. D. Monderer and D. Samet, Approximating common knowledge with common beliefs, *Games Econ. Behavior* **1** (1989), 170–190.
14. B. Skyrms, "Causal Necessity," Yale University Press, New Haven, CT, 1980.