Measurability Is Not About Information

By

Juan Dubra and Federico Echenique

March 2001

Measurability Is Not About Information [*]

Juan Dubra [†] and Federico Echenique [‡]

March 19, 2001

Abstract

We comment on the relation between models of information based on signals/partitions, and those based on σ-algebras. We show that more informative signals need not generate finer σ-algebras, hence that Blackwell's theorem fails if information is modeled as σ-algebras. The reason is that the σ-algebra generated by a partition does not contain all the events that can be known from the information provided by the signal. We also show that there is a non-conventional σ-algebra that can be associated to a signal which does preserve its information content. Further, expectations and conditional expectations may depend on the choice of σ-algebra that is associated to a signal. We provide a simple characterization of when the model is robust to changes in the σ-algebras.

[†]Corresponding author. Cowles Foundation, Yale University, PO Box 208268, New Haven, CT 06520-8268, and Universidad de Montevideo.email: `juan.dubra@yale.edu`

[‡]Department of Economics, UC Berkeley,and Departamento de Economía, Facultad de Ciencias Sociales, Universidad de la República, email: `fede@econ.berkeley.edu`.

§ 1 Introduction

In economics and game theory, information is usually modeled as signals on, or partitions of, a state space. For example, suppose a decision maker (DM) is evaluating the possibility of investing in a certain asset. There is a set Ω, the elements of which are interpreted as states of nature. The asset's rate of return is unknown to DM; but she has access to her favorite financial analyst's forecast, who announces a rate of return $Y(\omega)$ for each ω in Ω—we call the function Y a signal. For each announcement y, DM will only know that some ω with $y = Y(\omega)$ has occurred. Therefore, the signal Y generates a partition of Ω, which is the informational content of Y.

Still, it is very common in finance and statistics, but also in economics, to represent DM's information as a σ-algebra of subsets of Ω; this is because we need σ-algebras to use the machinery of probability theory. There are three different ways in which σ-algebras are used in association with information. First, given a signal Y, the smallest σ-algebra for which Y is measurable is often used (for example in the computation of conditional expectations). Second, the σ-algebra generated by Y's partition is used (as in Magill and Quinzii (1996)). Finally, sometimes a collection of σ-algebras is taken as the primitive set of information structures (Allen 1983, Allen 1986, Stinchcombe 1990). In all these cases, a σ-algebra is more informative than another if it contains more sets, i.e. if more sets are measurable.

Blackwell's (1951) theorem says that, given two signals Y and Z, Y is preferred to Z by every DM if and only if Y's partition is finer than Z's partition. We ask the following question: Given two signals Y and Z, is it true that Y is preferred to Z by every DM if and only if the σ-algebra of Y is finer than the σ-algebra of Z, for some natural definition of "the σ-algebra of?" In other words, do the σ-algebras usually associated to a signal preserve their informativeness? We provide the following results.

First, we show that Blackwell's theorem does not hold if information is modeled as σ-algebras in the usual way. Given one signal Y that is more informative than a signal Z, we show that the partition induced by Y does not, in general, generate a finer $\sigma-$algebra than the partition induced by Z. [1] We show that there is a loss of information when one goes from signals to the σ-algebras they generate. The reason for this loss is that if a DM's information allows her to decide whether any member of an arbitrary collection of sets is true or not, then she can decide whether their union is true or not. Since σ-algebras require that this be true only for countable collections of sets, the "knowledge" of sets that can only be built as arbitrary unions of elements in the partition is lost.

The loss of information in going from partitions to σ-algebras makes a difference in decision problems. We present a simple example where a DM must first choose a signal

---

[1]This statement is still true when we consider the smallest σ-algebras that make, respectively, Y and Z measurable, as well. This is discussed below.

and then, after observing its value, chooses to invest in a risky asset or in a riskless bond. If σ-algebras are used to model information DM chooses the least informative signal.

We show that the loss of information in passing from signals to σ-algebras can be avoided by associating a non-conventional σ-algebra to the signals. This is the σ-algebra of arbitrary unions of elements of the partition. Of course, this σ-algebra coincides with the σ-algebra generated by the signal's partition when the partition has a countable number of elements. Hence information is preserved exactly if, and only if, partitions have a countable number of elements (as in Magill and Quinzii (1996)). The information-preserving σ-algebra has the disadvantage that it has no clear links to the spaces' topological properties (like the Borel), and that it may be too large for some countably additive measures to be well-defined. Still, it is the only σ-algebra construction that preserves information—we show that Blackwell's theorem holds when the σ-algebra associated to signal is the information-preserving σ-algebra.

Second, we show that expectations and conditional expectations depend on the choice of σ-algebra that should be associated to a signal. This is important because, in the same model, different σ-algebras may yield different conclusions. We provide a simple characterization of when the model is robust to the choice of which σ-algebra should be associated to a signal.

Authors who use σ-algebras to model information do so for technical convenience, but they are apparently unaware that σ-algebras are unrelated to Blackwell-informativeness. For example, consider the following quote from Durrett (1996, p. 221), the leading graduate textbook in probability theory ($\mathcal{F}$ is a σ-algebra): "Intuitively we think of $\mathcal{F}$ as describing the information we have at our disposal—for each $A \in \mathcal{F}$, we know whether or not A has occurred." Or, in her seminal work on information in economics, Allen (1983, p. 66) "The abstract measure space $(\Omega, \mathcal{F})$ is arbitrary and can be interpreted to encompass notions of information partitions, signals, noisy observations, etc." The finance literature presents the same heuristics, "The filtration generated by a given stochastic process may be treated as a model of the information received through time by observing the process.", (Duffie 1988, p. 133). The heuristic is inaccurate because a σ-algebra is the collection of sets of which DM is able to make a probabilistic judgment, not the sets that DM can decide if they are true.

§ 2 Non-equivalence

This section is organized as follows. In § 2.0 we introduce some standard definitions and notation. In § 2.1 and § 2.3 we show that signals and σ-algebras are incompatible as models of information. In § 2.2 and § 2.4 we discuss a simple example where, if the σ-algebras approach is used, we predict that a DM will choose the least informative of two signals. In § 2.5 we explain why information is lost in the passage from signals to

σ-algebras, and we show that closedness with respect to arbitrary unions is necessary and sufficient to preserve information. In § 2.6 we argue that focusing on countable state-spaces may not be a solution in many interesting contexts.

§ **2.0**. A **partition** τ of a set Ω is a collection of pairwise disjoint subsets whose union is Ω; note that for each state of nature ω there is a unique element of τ that contains ω. A decision maker whose information is represented by τ is informed only that the element of τ that contains the true state of nature has occurred. In other words, the decision maker cannot distinguish between states that belong to the same element of τ. If Ω is a set, and $Y : \Omega \to \mathbf{R}$ is a function, define the partition P_Y of Ω associated to Y by

$$P_Y = \left\{ Y^{-1}(x) : x \in \mathbf{R} \right\}.$$

If τ, τ' are partitions, say that τ' is **finer** than τ if, for every $A \in \tau'$, there is B in τ such that $A \subseteq B$. We say that a signal Y is **more (Blackwell) informative** than a signal Z if P_Y is finer than P_Z.

A σ-**algebra** of subsets of a state-space Ω is a collection of sets that contains Ω and that is closed under countable unions and formation of complements. For any collection C of subsets of Ω, the σ-algebra **generated** by C, denoted $\sigma(C)$, is the smallest σ-algebra that contains C. A pair $(\Omega, \mathcal{F})$, where $\mathcal{F}$ is a σ-algebra of subsets of Ω, is a measurable space. Let $(\Omega, \mathcal{F})$ and $(\Theta, \mathcal{G})$ be two measurable spaces. A function $Y : \Omega \to \Theta$ is **measurable** if $Y^{-1}(B) \in \mathcal{F}$ for all $B \in \mathcal{G}$. The σ-algebra generated by Y, denoted $\sigma(Y)$, is the smallest σ-algebra for which Y is measurable.

A probability space is a triple $(\Omega, \mathcal{F}, P)$ where $\mathcal{F}$ is a σ-algebra of subsets of Ω and P is a probability measure. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$, the conditional expectation, $E(f|\mathcal{G})$, of any measurable function $f : \Omega \to \mathbf{R}$ is the equivalence class of $\mathcal{G}$-measurable functions $g : \Omega \to \mathbf{R}$ that satisfy $\int_A f dP = \int_A g dP$ for all $A \in \mathcal{G}$. Following standard practice, if Y is a signal we define $E(f|Y) = E(f|\sigma(Y))$ and $E(f|P_Y) = E(f|\sigma(P_Y))$.

The following lemma is useful, its proof follows directly from checking definitions, and therefore omitted.

Lemma A. *Let τ be a partition of Ω, then $\sigma(\tau)$ is the collection of sets A such that A or A^c is a countable union of elements of τ.*

The following observation (obvious in light of Lemma A) clarifies the interpretation of the σ-algebra generated by a partition.

Observation. *Let $(\Omega, \mathcal{F}, P)$ be a probability space. If π is a partition of Ω, all of whose elements are measurable, and such that $P(A) = 0$ for all $A \in \pi$; then $\sigma(\pi)$ is trivial, that is $P(A) \in \{0, 1\}$ for all $A \in \sigma(\pi)$.*

§ **2.1**. Let the state of the world be a number between 0 and 1, so that $\Omega = [0, 1]$, and suppose that a DM can observe one of two signals Y and Z about the state of the world.

Suppose that Y is given by $Y(\omega) = \omega$ for all ω in $[0, 1]$. The signal Z, in turn, is

$$Z(\omega) = \begin{cases} 0 & \text{if } \omega < \frac{1}{2} \\ 1 & \text{otherwise} \end{cases}$$

Of course, Y is more Blackwell informative than Z, Y's partition is $P_Y = \{\{\omega\} : 0 \le \omega \le 1\}$, Z's partition is $P_Z = \{[0, 1/2); [1/2, 1]\}$, and Y's partition is finer than Z's. By using Y DM would be able to distinguish between more states of nature. We shall show that the σ-algebras associated to P_Y and P_Z do not preserve this informational ranking.

Let us represent the informational content of signals Y and Z by σ-algebras. When partitions P_Y and P_Z are the primitive informational structure, the natural candidates are $\sigma(P_Y)$ and $\sigma(P_Z)$, the σ-algebras generated by P_Y and P_Z, respectively (Magill and Quinzii 1996). It is somewhat more frequent to use the σ-algebra generated by the signal viewed as a function, here $\sigma(Y)$ and $\sigma(Z)$ (see § 2.0). We show below that, in that case, problems in addition to those highlighted in this section appear. Further, some authors avoid uncountable partitions like P_Y, we discuss this in § 2.6.

Now, (by Lemma A in § 2.0), $\sigma(P_Y)$ is the collection of sets that are countable, or have a countable complement, while $\sigma(P_Z) = \{\emptyset; [0, 1/2); [1/2, 1]; \Omega\}$. Thus, $\sigma(P_Y)$ is not finer than $\sigma(P_Z)$ (in fact, $\sigma(P_Y)$ and $\sigma(P_Z)$ are incomparable), and Y is not more informative than Z when the σ-algebras model is used.

To the best of our knowledge, that the two notions of informativeness are incompatible is not known. To make this point, note that, for example, Yannelis (1991) claims that finer partitions generate finer sigma algebras.

§ **2.2** The following example shows that, in a standard decision problem, if σ-algebras are used to represent the informational content of two signals, DM will choose the least informative signal.

Let Ω, Y, and Z be as in § 2.1. Endow Ω with the Borel σ-algebra; $\sigma(P_Y)$ and $\sigma(P_Z)$ are sub-σ-algebras of the Borel. Let P be the Lebesgue measure on Ω.

Let X be the rate of return of some asset, suppose $X(\omega) = \omega$, and suppose there is a bond that yields a return of $3/8$ in every state of the world. Let the preferences of DM be such that she only cares about the expected return to her investment. After learning the value of her signal, she uses her prediction of X to decide whether to invest in the asset or in the bond.

We show below that $E(X|P_Y) = 1/2$ a.s. (specifically, that any version of $E(X|P_Y)$ equals a.s. the constant function $1/2$). It is immediate that

$$E(X|P_Z)(\omega) = \begin{cases} 1/4 & \text{if } \omega < 1/2 \\ 3/4 & \text{otherwise} \end{cases}$$

Now it is easy to check that DM will prefer signal Z over Y. In almost every state of

the world $E(X|P_Y)$ is larger than the return to the bond, so if DM has signal Y she will a.s. buy the asset. The ex-ante expected utility of choosing signal Y is thus $1/2$. On the other hand, if DM uses Z then she will buy the bond when $\omega < 1/2$ and the asset if $\omega \geq 1/2$, and will thus get a utility of $(3/8)(1/2) + (3/4)(1/2) = 9/16$. Hence, while Y is more informative than Z it is worth less to DM.

We now show that, indeed, $E(X|P_Y) = 1/2$ a.s., that is, that any version of $E(X|P_Y)$ is a.s. equal to $1/2$. Let $f : \Omega \to \mathbf{R}$ be any version of $E(X|P_Y)$. We shall show that $f^{-1}(1/2, +\infty)$ and $f^{-1}(-\infty, 1/2)$ are countable sets; this suffices as countable sets have probability zero. We first show that, for arbitrary natural n, $f^{-1}\left[\frac{1}{2} + \frac{1}{n}, +\infty\right)$ is countable. Suppose it is uncountable, then $(f^{-1}\left[\frac{1}{2} + \frac{1}{n}, +\infty\right))^c$ is countable, because f is $\sigma(P_Y)$-measurable and $\sigma(P_Y)$ is the countable-co-countable σ-algebra. Then $f^{-1}\left[\frac{1}{2} + \frac{1}{n}, +\infty\right)$ has measure one and we obtain,

$$\frac{1}{2} = \int\limits_{f^{-1}\left[\frac{1}{2}+\frac{1}{n}, +\infty\right)} f(\omega)dP(\omega) \geq \int\limits_{f^{-1}\left[\frac{1}{2}+\frac{1}{n}, +\infty\right)} \left(\frac{1}{2} + \frac{1}{n}\right) dP = \frac{1}{2} + \frac{1}{n},$$

a contradiction. That $f^{-1}\left(\frac{1}{2}, +\infty\right)$ is countable follows, as

$$f^{-1}(1/2, +\infty) = \bigcup_n f^{-1}[1/2 + 1/n, +\infty).$$

Similarly, $f^{-1}(-\infty, 1/2)$ is countable.

§ **2.3**. We can view signals Y and Z as random variables, and compare the generated σ-algebras. In the case of Y and Z from § 2.1, the σ-algebra generated by Y is indeed larger than the one generated by Z. This is not the general case, though.

Let $\mathcal{F}$ be any σ-algebra of subsets of $\mathbf{R}$ such that $\{1\} \in \mathcal{F} \neq 2^{\mathbf{R}}$. If $Y : \mathbf{R} \to \mathbf{R}$ is the identity, and $Z = \chi_E$ for some $E \in 2^{\mathbf{R}} \backslash \mathcal{F}$, we obtain $\sigma(Y) = \mathcal{F}$, and $\sigma(Z) = \{\emptyset, E, E^c, \mathbf{R}\}$. Thus $\sigma(Y)$ and $\sigma(Z)$ are not comparable, while Y is more Blackwell informative than Z. That is, Blackwell's theorem fails when the σ-algebra associated to the signals is $\sigma(Y)$ and $\sigma(Z)$, respectively.

Example B is a concrete version of the above: There is a measurable space $(\Omega, \mathcal{F})$ and two (Borel) signals $Y, Z : \Omega \to \mathbf{R}$ such that Y is more informative than Z, but $\sigma(Y)$ and $\sigma(Z)$ are incomparable. The measure space that we use was used in the 40's by Dieudonné, Doob, Kolmogorov to show that, if it is endowed with an extension of Lebesgue measure, conditional probabilities may not exist (Blackwell 1956).

Example B. *Let $\Omega = \mathbf{R}$, and let $E \subseteq \Omega$ have outer Lebesgue measure 1 and inner Lebesgue measure 0 (e.g. see the construction in Theorem D, in §13 Halmos (1950)). Let $\mathcal{F}$ on Ω be the σ-algebra generated by the Borel σ-algebra and E. Let $Y, Z : \Omega \to \mathbf{R}$ be $Z = \chi_E$, the indicator function of E, and Y be the identity, Y and Z are measurable. Then, $P_Z = \{E, E^c\}$, and $P_Y = \{\{\omega\} : \omega \in \Omega\}$—signal Y is better than signal Z, since*

6

the partition it generates is finer. Now, $\sigma(Z) = \{\emptyset, E, E^c, \Omega\}$, and $\sigma(Y)$ is the Borel σ-field on $\mathbf{R}$. Thus $\sigma(Z) \not\subseteq \sigma(Y)$ and $\sigma(Y) \not\subseteq \sigma(Z) : \sigma(Y)$ and $\sigma(Z)$ are not comparable.

We should stress that, while the interpretation of $\sigma(Y)$ as the information content of Y is emphasized in much of probability theory, we do not claim that the intuitions in the large body of literature on e.g. martingales is false. We simply wish to point out some inaccuracies in the heuristics behind the theory. In many common contexts no practical problems arise: If $\{X_n\}_{n=1}^\infty$ is a sequence of random variables on a probability space $(\Omega, \mathcal{F}, \mu)$, the increasing sequence $\{\sigma(X_1, \ldots, X_n)\}_{n=1}^\infty$ is said to represent the "unfolding of information". While the examples above show that $\sigma(X_1, \ldots X_n)$ should not be called "the information in periods $1, \ldots n$", it is clear that information is increasing over time (the partitions associated to $(X_1, \ldots X_n)$ are finer for larger n).

Nevertheless, in interpreting $\sigma(Y)$ as information there are problems in addition to the issues raised in § 2.1 (and discussed further in § 2.5). It is immediate to see the problem with interpreting $\sigma(Y)$ as informational content of Y: $\sigma(Y)$ is the collection of sets whose probability we can determine from the distribution of Y; on the other hand, its associated partition P_Y is the collection of sets A such that we can decide if A is true or not from the value of Y. Moreover, the partition P_Y, Y's informational content, is *completely* unrelated to the σ-algebra on the set where Y takes its values, while the σ-algebra generated by Y is *exactly* the set of preimages through Y of all measurable sets.

This point is clearer if we introduce a probability measure on $[0, 1]$, let P be the Lebesgue probability measure. Let Y and Z be as in § 2.1. Signal Y is such that $\sigma(P_Y)$ is trivial, while $\sigma(Y)$ is the Borel σ-algebra, and thus for any $\lambda \in (0, 1)$ there is $E \in \sigma(Y)$ with $P(E) = \lambda$. The first reflects that Y will show than some element $A \in P_Y$ occurred, which has probability zero, and its complement A^c has probability one. On the other hand, $\sigma(Y)$ reflects that the σ-algebra on the target space is so rich that one can obtain sets of any measure, something quite unrelated to Y's informational content.

§ 2.4 The counter-intuitive prediction in § 2.2 obtains also when the informational content of Y and Z is modeled by $\sigma(Y)$ and $\sigma(Z)$. The source of the anomaly in § 2.3 is that $E(X|P_Y) = 1/2$. Is this also true when we use $\sigma(Y)$ to predict X? The answer is that it depends on the σ-algebra on Y's target space. If $\mathbf{R}$ is endowed the Borel σ-algebra then the identity is a version of $E(X|\sigma(Y))$, but if $\mathbf{R}$ is endowed with the σ-algebra of countable sets or complements of countable sets, then every version of $E(X|\sigma(Y))$ must equal $1/2$ a.s.

Besides reiterating the point in § 2.3, this shows that, with probability 1, $E(X|\sigma(Y))$ may be different if different σ-algebras are imposed on Y's target space. By changing the σ-algebra on the signals' target space we reverse DM's preferences among the available signals. Moreover, the σ-algebra on the target space does not have, in our opinion, a natural interpretation. If we view this σ-algebra as a model of DM's ability to "perceive"

the values of the signals, then we run into the same problems as when we use σ-algebras to model information on Ω. In particular, the ability to perceive is not closed under arbitrary unions, see § 2.5.

§ **2.5**. The problem presented in § 2.1 and § 2.3 has its origin in the construction of the σ-algebras associated to a signal. If we instead construct σ-algebras by an operation that exactly preserves the informational content of the signal, the problem does not arise, and Blackwell's theorem is recovered. For a signal Y, let σ_Y be the collection of sets which are arbitrary unions of elements in the partition P_Y. It is immediate to check that σ_Y is a σ-algebra.

Lemma C. *Let Ω and Θ be two sets, and Y, Z two functions from Ω to Θ. Y is more informative than Z if and only if $\sigma_Z \subseteq \sigma_Y$.*

Proof. If Y is more informative than Z, for every collection of sets $\{A_i\}$ in P_Z, there exists a collection $\{B_i\}$ in P_Y such that $\bigcup_i A_i = \bigcup_i B_i$, so $\sigma_Z \subseteq \sigma_Y$. Suppose now that $\sigma_Z \subseteq \sigma_Y$, but that Y is not more informative than Z. This means that there exists a set A in P_Z such that for every collection $\{B_i\}$ in P_Y, $A \neq \bigcup_i B_i$. Thus, $A \in \sigma_Z$, but $A \notin \sigma_Y$, a contradiction. ∎

Lemma C and Blackwell's theorem imply then,

Theorem D. *Let Ω and Θ be two sets, and Y, Z two functions from Ω to Θ. Every DM prefers Y over Z if and only if $\sigma_Z \subseteq \sigma_Y$.*

Why does the σ_Y construction preserve information? If a DM's information allows her to decide if a member of an arbitrary collection of sets is true or not, then she can decide if their union is true or not. Thus σ_Y is obtained by adding all those sets that DM "knows". There is a problem in §2.1 because the σ-algebra generated by P_Y does not contain all sets that are uncountable unions of elements of P_Y. Now, when Ω is finite or countably infinite σ_Y and $\sigma(P_Y)$ coincide, and more information translates into larger σ-algebras. [2]

Allen (1983) (see also Stinchcombe (1990)) presents a model of information where the set of primitive information structures is a collection of σ-algebras on a state space. Our observations imply that, in order for a collection of σ-algebras to be interpreted as information structures, they should be closed under arbitrary unions. This remark is subject to the following qualification: Stinchcombe (1990) proves that, in certain well-behaved spaces, a countably generated σ-algebra $\mathcal{F}$ can be identified with the partition $\{\cap \{B : B \in \mathcal{F}, \omega \in B\} : \omega \in \Omega\}$ of its *atoms*, and that all σ-algebras are "close" to a countably generated σ-algebra. [3] In this particular sense, then, arbitrary σ-algebras possess an informational content.

---

[2]This explains why authors that use finite state spaces, like Magill and Quinzii (1996), have a consistent model of information, even though they go back and forth between partitions and σ-algebras.

[3]We are grateful to Maxwell Stinchcombe for pointing this out.

§ 2.6 Many authors use a countable state space, or, with an uncountable state space, focus only on finite or countable partitions. This practice avoids the problems that we have pointed out, and probably people follow this practice because they are aware that there are problems associated to uncountable partitions. To the best of our knowledge, no one has identified the problems we point to, and their link to the modeling of information. In fact, people normally rule out uncountable partitions without providing a justification for doing so.

Many models require an uncountable number of states of nature. This is the case, for example, of Savage's model of decision under uncertainty (Fishburn 1970) or of games of incomplete information (Mertens and Zamir 1985, Brandenburger and Dekel 1993). In other models, an uncountable space is necessary to use calculus methods. If we wish to model information as partitions in a model with an uncountable state space we cannot restrict to countable partitions—we would really be considering a countable state space by ruling out, among other things, that agents distinguish among an uncountable number of states.

§ 3 Some additional problems and a fix

Let Y be a signal. If X is a random variable, we show in § 3.1 that $E(X|Y)$ and $E(X|P_Y)$ may differ, and in § 3.2 that $E(X)$ may depend on whether $\sigma(Y)$ or $\sigma(P_Y)$ is used. The source of this problem is, of course, that $\sigma(Y)$ and $\sigma(P_Y)$ differ; in § 3.3 we characterize the cases when they do not differ, and show that no problem arises when partitions are taken as primitives throughout.

§ 3.1. Let $\Omega = [0,1]$, endowed with the Borel σ-algebra, and $Y : \Omega \to \Omega$ be the identity. Then, $\sigma(Y)$ is the Borel σ-algebra on Ω, and $\sigma(P_Y)$ is the collection of $A \subseteq \Omega$ such that A or A^c is countable. Let $X : \Omega \to \mathbf{R}$ be one-to-one, then X is a version of $E(X|Y)$ but not of $E(X|P_Y)$, as X is not $\sigma(P_Y)$-measurable.

This example suggests that we may need to make a choice between $E(X|Y)$ and $E(X|P_Y)$ when modeling DM's prediction of X based on the information in Y.

§ 3.2. Let Ω and Y be as in § 3.1, and let X be the identity too. Let P be the Lebesgue (uniform) measure on $(\Omega, \sigma(Y))$, and on $(\Omega, \sigma(P_Y))$. We now show that $E(X)$ depends on whether Ω is endowed with $\sigma(Y)$ or $\sigma(P_Y)$. Denote by $\Psi_{\mathcal{F}}$ the collection of simple functions when Ω is endowed with the σ-algebra $\mathcal{F}$. Then,

$$\sup_{\left\{\psi \in \Psi_{\sigma(P_Y)} : \psi \leq X\right\}} \int \psi(\omega) dP(\omega) = 0 < 1/2 = \sup_{\left\{\psi \in \Psi_{\sigma(Y)} : \psi \leq X\right\}} \int \psi(\omega) dP(\omega).$$

Hence $E(X) = 0$ in $(\Omega, \sigma(P_Y), P)$, and $E(X) = 1/2$ in $(\Omega, \sigma(Y), P)$.

9

§ 3.3. The source of these problems is the difference between $\sigma(Y)$ and $\sigma(P_Y)$. We can fix these problems by taking partitions as basic throughout. If $Y : \Omega \to \Theta$ is a signal, and π a partition of Θ, we define P_Y^π to be the partition formed by the preimages through Y of the elements of π; σ_Y is, as before, the set of unions of elements in P_Y^π.

Proposition E. *Let $Y : \Omega \to \Theta$ and π be a partition of Θ. If Θ is endowed with $\sigma(\pi)$, $\sigma(P_Y^\pi) = \sigma(Y)$. If Θ is endowed with the σ-algebra of unions of elements of π, then $\sigma_Y = \sigma(Y)$.*

By different partitions on Θ we can study different degrees of "perceptiveness", i.e. of the quality of information in the signal's target space. We always assume, though, that π consists of all singletons; in this case we can characterize fully the situations where $\sigma(P_Y) = \sigma(Y)$.

Proposition F. *Let $(\Omega, \mathcal{F})$ and $(\Theta, \mathcal{G})$ be two measurable spaces, and let $Y : \Omega \to \Theta$ be onto. $\mathcal{G}$ is generated by the singletons in Θ if and only if $\sigma(P_Y) = \sigma(Y)$. $\mathcal{G} = 2^\Theta$ if and only if $\sigma_Y = \sigma(Y)$.*

We omit the proof of Proposition E because its proof is similar to the proofs of the *only if* parts of Proposition F.

Proof of Proposition F: We shall first prove the first equivalence. Let π be the collection of all singletons in Θ, and $\mathcal{G} = \sigma(\pi)$. Then, for all $B \in P_Y$, there is $\{\theta\} \in \pi$ such that $B = Y^{-1}(\theta)$. Since $\{\theta\} \in \sigma(\pi)$ we obtain $B \in \sigma(Y)$; hence $P_Y \subseteq \sigma(Y)$, and $\sigma(P_Y) \subseteq \sigma(Y)$. Let $B \in \sigma(Y)$, then there is $A \in \sigma(\pi)$ with $B = Y^{-1}(A)$. If $A = \cup\{A_i : i = 1, 2, \ldots\}$ with $A_i \in \pi$ for $i = 1, 2, \ldots$, then $B = Y^{-1}(A) = \cup\{Y^{-1}(A_i) : i = 1, 2, \ldots\}$. Then, $B \in \sigma(P_Y)$, as $Y^{-1}(A_i) \in \sigma(P_Y)$ for $i = 1, 2, \ldots$.

By Lemma A, if A is not a countable collection of elements of π, A^c is, and thus $B^c \in \sigma(P_Y)$, so $B \in \sigma(P_Y)$. Hence $\sigma(P_Y) = \sigma(Y)$.

Suppose now that $\sigma(P_Y) = \sigma(Y)$. For all $A \in \mathcal{G}$, $B = Y^{-1}(A)$ belongs to $\sigma(P_Y)$. Suppose B is the countable union of elements of P_Y. Then, $Y(B) = A$ is the countable union of singletons in π, and thus $A \in \sigma(\pi)$. If B is uncountable, B^c is countable and $Y(B^c) = A^c$ is the countable union of singletons in π, and thus $A^c, A \in \sigma(\pi)$. This shows that $\mathcal{G} \subseteq \sigma(\pi)$.

To show that $\sigma(\pi) \subseteq \mathcal{G}$, we will show that for all $\theta \in \Theta$, $\{\theta\} \in \mathcal{G}$. Let $B = Y^{-1}(\theta) \in P_Y \in \sigma(P_Y) = \sigma(Y)$. Since $\sigma(Y)$ is the σ-algebra of all the preimages of all $\mathcal{G}$-measurable sets, there is some $A \in \mathcal{G}$ such that $Y(B) = A$. But then $A = \{\theta\}$.

To prove the second statement let $\mathcal{G} = 2^\Theta$ and let $A \in \sigma(Y)$. Then $A = Y^{-1}(B) = \cup\{Y^{-1}(\theta) : \theta \in B\}$ for some $B \in \mathcal{G}$. Now, $Y^{-1}(\theta) \in P_Y$ for every $\theta \in B$, and thus $A \in \sigma_Y$ by definition of σ_Y. Hence, $\sigma(Y) \subseteq \sigma_Y$. To prove $\sigma_Y \subseteq \sigma(Y)$, let $A \in \sigma_Y$. Then

10

$A = Y^{-1}(B) = \cup\{Y^{-1}(\theta) : \theta \in B\}$, for some $B \subseteq \Theta$. Since $B \in \mathcal{G}$, $A \in \sigma(Y)$. Thus $\mathcal{G} = 2^\Omega$ implies that $\sigma_Y = \sigma(Y)$.

Let $\sigma_Y = \sigma(Y)$ and $B \subseteq \Theta$. Then, $Y^{-1}(B) = \cup\{Y^{-1}(\theta) : \theta \in B\} \in \sigma_Y$, by definition of σ_Y. Then $Y^{-1}(B) \in \sigma(Y)$, so there is $C \in \mathcal{G}$ with $Y^{-1}(C) = Y^{-1}(B)$, but then $B = C$ as Y is onto. Thus $B \in \mathcal{G}$. ∎

References

ALLEN, B. (1983): "Neighboring Information and Distributions of Agents' Characteristics Under Uncertainty," *Journal of Mathematical Economics*, 12, 63–101.

——— (1986): "The demand for (differentiated) information," *The Review of Economic Studies*, 53(3), 311–323.

BLACKWELL, D. (1951): "Comparison of Experiments," in *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, ed. by J. Neyman, pp. 93–102. University of California Press, Berkeley and Los Angeles.

——— (1956): "On a Class of Probability Spaces," in *Proceedings of the third Berkeley symposium on mathematical statistics and probability*, ed. by J. Neyman, pp. 1–6. University of California Press, Berkeley and Los Angeles.

BRANDENBURGER, A., AND E. DEKEL (1993): "Hierarchies of Beliefs and Common Knowledge," *Journal of Economic Theory*, 59(1), 189–98.

DUFFIE, D. (1988): *Security Markets: Stochastic Models*. Academic Press.

DURRETT, R. (1996): *Probability: Theory and Examples*. Duxury Press.

FISHBURN, P. C. (1970): *Utility Theory for Decision Making*. Wiley.

HALMOS, P. R. (1950): *Measure Theory*. Van Nostrand (Reprinted by Springer Verlag, 1974).

MAGILL, M., AND M. QUINZII (1996): *Theory of Incomplete Markets*. MIT Press.

MERTENS, J.-F., AND S. ZAMIR (1985): "Formulation of Bayesian Analysis for Games of Incomplete Information," *International Journal of Game Theory*, 14, 1–29.

STINCHCOMBE, M. B. (1990): "Bayesian Information Topologies," *Journal of Mathematical Economics*, 19, 233–253.

YANNELIS, N. C. (1991): "The Core of an Economy with Differential Information," *Economic Theory*, 1(2), 183–97.