The Annals of Statistics
1981, Vol. 9, No. 1, 59–65

LEARNING AND DECISION MAKING WHEN SUBJECTIVE PROBABILITIES HAVE SUBJECTIVE DOMAINS

BY CHARLES F. MANSKI

The Hebrew University of Jerusalem

This paper relaxes the conventional subjective probability setup by allowing the σ-algebra on which probabilities are defined to be subjective along with the probability measure. First, the role of the probability domain in existing statistical decision theory is examined. Then the existing theory is extended by characterizing the individual's selection of a probability domain as the outcome of a decision process.

1. Introduction. A central element of modern statistical decision theory is the idea of a subjective probability system (Θ, Ω, P), that is a probability measure P subjectively imposed on a σ-algebra Ω of subsets of an abstract space Θ. Within the existing theory the probability domain (Θ, Ω) is prespecified and fixed. The measure P on the other hand is variable. Not only may P vary across individuals but a given individual's subjective probabilities can vary as a function of the data available to him. Bayes rule, of course, provides the mechanism by which data is integrated with prior beliefs. This paper relaxes the conventional subjective probability setup by allowing the σ-algebra on which probabilities are defined to be variable along with the probability measure.

The plan of the paper is as follows. Section 2 reviews elements of existing statistical decision theory, in particular the conditions a probability domain must satisfy if Bayes rule and expected utility maximization are to be well-defined processes. Section 3 presents a model of decision making when the elicitation of subjective probabilities is costly. In this context, the decision maker faces the auxiliary problem of selecting a probability domain.

Before proceeding it must be said that the present paper does not stand in isolation from past literature. The Keynes (1921) distinction between risk and uncertainty, the Good (1962) discussion of the 'measure of a nonmeasurable set', the Fine (1973) suggestion that probabilities need not be defined on a σ-algebra and the extensive Dempster (1967, 1968)-Shafer (1976) work on inference based on 'lower probabilities' are all relevant and I have benefitted from the thinking of these authors.

2. Statistical decision theory with the probability domain fixed.

A. LEARNING.

1. *Conditions permitting Bayesian learning.* Given a subjective probability system (Θ, Ω, P), the conditions under which Bayesian learning is a well-defined process are most often assumed without comment. For present purposes, however, it is important to review and interpret these conditions.

Let (Y, Φ, ν) be a measure space of observations and F_θ, $\theta \in \Theta$ a family of probability measures on (Y, Φ), each $F_\theta \ll \nu$, with ν σ-finite. Let $f(y/\theta)$ be the density at $y \in Y$ of F_θ. Furthermore, let μ be a measure on (Θ, Ω) with $P \ll \mu$ and μ σ-finite and let $p(\theta)$ be the prior density at $\theta \in \Theta$. Finally, assume that for each $y \in Y$, the likelihood function $f(y/\cdot)$ is P-integrable and that $\int_\Theta f(y/\theta)\, dP(\theta) = 0 \Rightarrow f(y/\theta) = 0$, all $\theta \in \Theta$. Then Bayes rule defines a

Received March 1979; revised August 1979.

AMS 1970 *subject classification.* Primary 62A15, 62C10.

Key words and phrases. Subjective probability systems, Bayes rule, measurable utility, elicitation of subjective probabilities, statistical decision theory.

posterior density on Θ as follows. If (y, θ) satisfy $f(y/\theta)p(\theta) > 0$, then

$$p(\theta/y) = \frac{f(y/\theta)p(\theta)}{\int_\Theta f(y/\theta)\, dP(\theta)}.$$

Otherwise, $p(\theta/y) = 0$.

In the above, two conditions involve the subjective probability domain Ω. First, the P-dominating, σ-finite measure μ has domain Ω. Second, integrability of the likelihood function requires that this function be Ω-measurable.

Of the two conditions, the first is relatively inconsequential. The specification of Ω does delimit the class of measures μ which are σ-finite. However, given any Ω, the set of P-dominating σ-finite measures always is nonempty. In particular, $\mu = P$ is always a possible choice. Since the posterior probability on any $\omega \in \Omega$ does not depend on what P-dominating, σ-finite μ one employs, it follows that the restrictions Ω places on the class of feasible μ measures have no operational effect.

The second, measurability, condition is of substantial importance. An essential aspect of the restriction Ω imposes on $f(y/\cdot)$ is contained in the following simple proposition, whose validity is a direct consequence of the definition of measurability.

Proposition 1. *Let ω be a minimal set of Ω, that is $\omega \in \Omega$ and no nonempty proper subset of ω belongs to Ω. Then $f(y/\cdot)$ can be Ω-measurable only if there exists a $J_\omega \geq 0$ such that $f(y/\theta) = J_\omega$, all $\theta \in \omega$.*

It follows from the proposition that if the likelihood function is not constant on minimal sets of the probability domain, then Bayes rule is not well defined.

The Bayesian literature has generally ignored the implications of Proposition 1. This has been done without technical error because authors have universally specified probability systems (Θ, Ω, P) in which the minimal sets of Ω are the elements of Θ. In particular, in the case of Θ countable, it has generally been assumed that $\Omega = 2^\Theta$. In the case of $\Theta = R^N$, Θ has been taken to be the Borel σ-algebra or the Lebesgue measurable sets.[1]

What must be emphasized here is that the assumption of such rich probability domains buys integrability of the likelihood function at a price in behavioral realism. As an example, consider a simple economic growth model $e^{\theta T}$ where T is 'time' and $\theta \in \Theta = (-\infty, \infty)$. An economist may reasonably be willing to place subjective probability on the possibilities $\theta \in (-\infty, 0)$, $\theta = 0$, $\theta \in (0, \infty)$ corresponding to a shrinking, stable or expanding economy but to go no further. In this case, $\Omega = \sigma[\phi, (-\infty, 0), [0], (0, \infty)]$ where $\sigma[\cdot]$ denotes the operation of forming the smallest σ-algebra containing the sets in the brackets. Given an observation y, Bayesian learning about the economy's growth rate is then possible only if $f(y/\cdot)$ is constant over each of the sets $(-\infty, 0)$ and $(0, \infty)$, a strong restriction on the nature of the observation.

It is of interest to observe that by Proposition 1, Bayesian learning may not be possible because an observation is too informative, in the sense that its likelihood varies too much across the elements of Θ. Observations with sharply varying likelihood functions are usually and quite naturally considered to be the ones which carry the most information about θ. It is therefore curious that it is in the presence of such observations that failure of Bayes rule is most likely.

2. *Learning without Bayes rule.* When Bayes rule fails, how may one proceed?

One approach is simply not to let the rule fail. A way to accomplish this is to condition specification of the prior probability system on the observation. That is, given an observation

[1] It should be noted that identification of the minimal sets of Ω with the elements of Θ does not per se ensure that Ω is an interesting σ-algebra. For example, consider the case where $\Theta = R^1$ and Ω is generated from the points of R^1 plus R^1 itself. Then the only uncountable subset of R^1 which is a member of Ω is the set R^1 itself.

with associated likelihood function, one might impose a prior probability domain rich enough to support application of Bayes rule. The use of such a data-dependent prior domain may not be orthodox Bayesian procedure but it is akin in spirit to suggestions made by Zellner (1977) and others for the specification of prior probability measures.[1]

Alternatively, one might condition specification of the likelihood function on the prior probability domain. That is, given Ω, one might specify $f(y/\cdot)$ to be Ω-measurable. This procedure will not be appealing to those Bayesians who view the likelihood function as an objective characterization of a random process but may be acceptable to those who argue that the likelihood function is no less subjective than is the prior.

Of course mixtures of the above procedures are also possible. In fact, to those authors who take as their primitive the subjective product probability space $(\Theta \times Y, \Omega \times \Phi, F_\theta \cdot P)$ and derive the posterior via Bayes theorem, there exist no grounds for treating the prior and the likelihood asymmetrically.[2]

B. DECISION MAKING.

1. *Conditions permitting expected utility maximization.* Let D be a set of feasible actions and $U(d/\theta)$ be the utility associated with action $d \in D$ under state of nature $\theta \in \Theta$. Then the expected utility maximization decision rule directs one to select an action $d \in D$ that maximizes the expected utility $V(d) = \int_\Theta U(d/\theta)\, dP(\theta)$.

The expected utility criterion is well defined only if for each $d \in D$, $U(d/\cdot)$ is P-integrable and if the function $V(d)$ has a maximum in D. The probability domain Ω is involved in these conditions in that $U(d/\cdot)$ must be Ω-measurable if it is to be P-integrable.

Recognition of this requirement is long standing. Nevertheless, the measurability condition on U, like that on the likelihood function, has been largely ignored in the recent literature. Again, the assumption by authors of rich σ-algebras has made this lack of attention technically legitimate. But again, a price in behavioral realism has been paid. In particular, Proposition 1 has the same force with respect to the utility function $U(d/\cdot)$ as it does with the likelihood function f.

2. *Other decision rules.* As in the case of Bayesian learning, when the utility function is not measurable, one may decide to respecify U, Ω, or both. The maximin and maximax decision rules may usefully be viewed as the results of two possible respecifications of U undertaken when $\Omega = (\phi, \Theta)$. In the maximin case, $U(d/\cdot)$ is replaced by $\underline{U}(d/\Theta) = \inf_{\theta\in\Theta} U(d/\theta)$, while the maximax replacement is $\bar{U}(d/\Theta) = \sup_{\theta\in\Theta} U(d/\theta)$. The functions $\underline{U}(d/\Theta)$ and $\bar{U}(d/\Theta)$, each being constant over Θ, are both measurable under Ω.

3. Statistical decision theory with the probability domain variable. Existing literature implicitly adopts the position that probability domains, like utilities and prior subjective probability measures, are idiosyncratic personal characteristics which should be taken as exogenous by decision theory.

Alternatively, one can regard a subjective probability domain as a construct selected by an individual for its value in a given context. In this view, a domain is not an innate attribute of a person but rather a tool he employs in solving a decision problem. As the decision context changes, for example through the accumulation of observational evidence or a change in the set of feasible actions, the individual suitably alters the domain he uses. If this endogenous

[1] One major difference between the specification of data-dependent prior domains and data-dependent prior measures should be noted. In the latter case, the existence of extensive data will make the choice of prior measure largely irrelevant as the data will dominate the prior in forming the posterior. In the former case, however, extensiveness of the data does not reduce the importance of the prior specification. In application of Bayes rule, the posterior domain is always the same as the prior domain.

[2] Bayes theorem, that is the probability theory result that a measure on a product space can be decomposed into conditional and marginal measures, is to be distinguished from Bayes rule, where the prior (a marginal measure) and likelihood (a family of conditional measures) are posited as primitives. The presentation of Bayesian analysis in DeGroot (1970) utilizes Bayes theorem. That in Lindley (1971) rests on Bayes rule.

view is adopted, one must seek to explain how probability domains are determined.

The latter position will be taken here. In particular, we shall adopt the premise that a person could in principle place a subjective probability measure on an arbitrarily refined system of subsets of the model space. However, because the elicitation of subjective probabilities is costly, he may choose to assign probabilities only on a limited domain. The reason why the person bothers to elicit his probabilities at all in the face of elicitation costs is that knowledge of the probabilities is useful in solving a statistical decision problem.

In order to move the discussion beyond generalities, I sketch out a formal model.

A. THE IDEALIZED DECISION PROBLEM. To begin, assume a decision maker with a set of feasible actions D, a model space Θ and utilities $U(d/\theta)$, $d \epsilon D$, $\theta \epsilon \Theta$. Assume that the utilities are bounded in $D \times \Theta$. Also assume that there exists a σ-algebra Ω of subsets of Θ such that $U(d/\cdot)$ is Ω-measurable for all $d \in D$. Moreover, if the elicitation of subjective probabilities were costless, the decision maker would place a subjective probability measure P on Ω.

These assumptions define an idealized decision problem whose solution is a $d^* \in D$ which maximizes the expected utility function $V(d) = \int_\Theta U(d/\theta)\, dP(\theta)$ over $d \in D$. The decision problem is idealized because the process of finding its solution has been assumed costless. In what follows, this assumption is relaxed.

B. ELICITATION COSTS. There can be no doubt that the process of solving a formal statistical decision problem can be costly to the problem solver. In particular, there is ample evidence that the introspective elicitation of subjective probabilities constitutes a difficult, time consuming task even in relatively simple contexts.[1]

To characterize the costs of eliciting subjective probabilities, assume that the decision maker has an initial domain Ω_0 on which he has previously assigned probabilities. Also assume that the decision maker's new domain Ω_1 will always be a refinement of the initial one. That is, $\Omega_0 \subset \Omega_1$. Now define $C(\Omega_1 \mid \Omega_0)$ to be the cost of eliciting probabilities on the refined domain Ω_1 given that the decision maker had previously elicited probabilities on the cruder domain Ω_0.

We shall assume that the cost function satisfies certain intuitive but nevertheless restrictive conditions. Let $\Omega_0 \subset \Omega_\alpha \subset \Omega_1$ with $\Omega_0 \neq \Omega_\alpha \neq \Omega_1$. Then we assume

(i) $\inf_{\Omega_1} C(\Omega_1/\Omega_0) = K > 0$
(ii) $C(\Omega_1/\Omega_\alpha) + C(\Omega_\alpha/\Omega_0) = C(\Omega_1/\Omega_0)$.

That is, elicitation costs are bounded below and path independent.

We shall also assume that probabilities elicited on crude and refined domains agree, that is

(iii) $$\omega \in \Omega_0, \Omega_1 \Rightarrow P_0(\omega) = P_1(\omega)$$

where P_0 and P_1 are the subjective probability measures on Ω_0 and Ω_1 respectively.

Finally, we assume that the elicitation cost $C(\Omega_1/\Omega_0)$ is known to the decision maker and is subtracted from the idealized utility $U(d/\theta)$ to form the joint utility of an action and a probability elicitation

(iv) $Q(d, \Omega_1/\theta, \Omega_0) = U(d/\theta) - C(\Omega_1/\Omega_0)$

C. THE DECISION PROCESS. Consider now a decision maker with initial probability domain Ω_0 such that the utilities $U(d/\cdot)$ are not Ω_0-measurable. The existence of elicitation costs faces this decision maker with a situation in which solution of the idealized decision problem is costly. And it poses for him the auxiliary problem of selecting a domain alongside his need to choose an action from D.

Let d_0 be the action the person would select conditional on Ω_0 being his domain. Rather

[1] The elicitation of utilities and computation of expected utilities may also be burdensome. In order to focus attention on the problem of selecting a probability domain however, I shall continue to make the idealized assumption that these latter aspects of decision making are costless.

than retaining Ω_0, the decision maker may elect to refine his probability domain before selecting an action from D. If Ω_1 is the refined domain and d_1 the selected action under Ω_1, the realized value of the probability elicitation is $V(d_1) - V(d_0)$ and its cost is $C(\Omega_1/\Omega_0)$. At the time of the elicitation decision, the elicitation cost is known to the decision maker but its value cannot be computed.

The above describes the context in which the decision maker finds himself. It seems useless in this context to seek an optimal solution to the decision maker's problem. As Simon (1957) properly points out, when the process of solving an idealized optimization problem is costly, the process of solving the respecified optimization problem which makes these costs explicit will generally be even more costly. I shall therefore suggest a decision process that a person facing elicitation costs might reasonably follow.

To introduce the suggested procedure, first observe that under conditions (i) and (ii) the cost of eliciting probabilities on an infinite system of sets is itself infinite.[1] It may therefore be assumed that the initial probability domain and subsequent refinements are all finite.

Next note that with any finite algebra of sets Ω_0 there is associated a unique basis Γ_0, a finite system of sets which partitions Θ into mutually exclusive and exhaustive subsets and which satisfies the condition $\sigma(\Gamma_0) = \Omega_0$. The sets $\gamma \in \Gamma_0$ are the minimal sets of Ω_0.

Third, for any $d \in D$, $\gamma \subset \Theta$, define the lower and upper utilities

$$\underline{U}(d/\gamma) = \inf_{\theta\in\gamma} U(d/\theta),\ \bar{U}(d/\gamma) = \sup_{\theta\in\gamma} U(d/\theta)$$

and thence the lower and upper expected utilities

$$\underline{V}_0(d) = \textstyle\sum_{\gamma\in\Gamma_0} \underline{U}(d/\gamma)\cdot P_0(\gamma),\ \bar{V}_0(d) = \textstyle\sum_{\gamma\in\Gamma_0} \bar{U}(d/\gamma)\cdot P_0(\gamma).$$

Observe that the lower and upper expectations are computable by a decision maker using domain Ω_0.

As a final preliminary, consider an action $\underline{d}_0 \in D$ which maximizes $\underline{V}_0(\cdot)$ and a $\bar{d}_0 \in D$ which maximizes $\bar{V}_0(\cdot)$. Given Ω_0, $\underline{d}_0$ and $\bar{d}_0$ are respectively the maximin and maximax solutions to the decision maker's choice problem. Recalling that U is assumed bounded, it is easy to see that

$$\infty > \bar{V}_0(\bar{d}_0) \geq \bar{V}_0(d^*) \geq V(d^*) \geq V(\underline{d}_0) \geq \underline{V}_0(\underline{d}_0) > -\infty.$$

We now assume that given Ω_0, the decision maker begins by solving the costless problem max $\underline{V}_0(d)$ as an approximation to the idealized problem max $V(d)$. With $\underline{d}_0$ the tentatively chosen action, the maximal value of refining the probability domain is $V(d^*) - V(\underline{d}_0)$. This uncomputable quantity is bounded from above by the computable expression $\bar{V}_0(\bar{d}_0) - \underline{V}_0(\underline{d}_0)$. It immediately follows that if there exists no refinement Ω_1 of Ω_0 such that $\bar{V}_0(\bar{d}_0) - \underline{V}_0(\underline{d}_0) > C(\Omega_1/\Omega_0)$, the decision maker should elect to retain the domain Ω_0 and take the action $\underline{d}_0$. It remains then to discuss situations in which a beneficial refinement may exist.

There are many a priori sensible strategies a decision maker might follow in this context and it is difficult to point to one as being clearly superior. One general principle, however, is that refinement of the domain should be approached as a sequential process rather than as a once and for all decision.

The value of a sequential strategy derives from the assumption that elicitation costs are path independent. Under path independence, the cost of shifting from a domain Ω_0 to a proposed refinement Ω_1 is independent of the number of intermediate refinements made en route. On the other hand, intermediate refinements yield probability information that allow the decision maker to reevaluate the transition to Ω_1 before incurring all of its costs. It follows that the best refinement processes are those with a maximal number of stages, that is, those in which each stage makes a binary division of one of the sets forming the current probability domain basis.

[1] To see this, consider $C(\Omega_0/(\phi, \Theta))$ where Ω_0 is infinite. Let Ω_{-1} be an infinite σ-algebra with $\Omega_{-1} \subset \Omega_0$, $\Omega_{-1} \neq \Omega_0$. Then $C(\Omega_0/(\phi, \Theta)) \geq C(\Omega_{-1}/(\phi, \Theta)) + K$. By induction, $C(\Omega_0/(\phi, \Theta))$ is greater than any positive number.

Given the above, the decision maker's elicitation problem reduces from one of considering all possible refinements of the initial domain Ω_0 to one of examining the minimal sets $\gamma \in \Gamma_0$ and determining which one if any of these sets should be split into two parts. Although a formal analysis of this problem will not be attempted here, attention may be called to three germane factors.

First, the relative values of the prior probabilities $P_0(\gamma)$, $\gamma \in \Gamma$ will be relevant. All else equal, the higher the probability assigned to a set, the more this set influences the expected utility values, hence the more important it is to know how the probability $P_0(\gamma)$ distributes itself within γ.

Second, one should be concerned with the relative spread of utility values within each set of Γ_0. If for some $\gamma \in \Gamma_0$ it is the case that $U(d/\theta) = U(d/\theta')$, all $\theta, \theta' \in \gamma$, $d \in D$ or if $U(d/\theta) = U(d'/\theta)$, all $d, d' \in D$, $\theta \in \gamma$ then clearly there is no value in refining γ. Conversely, refinement has the greatest potential benefit when the utilities $U(d/\theta)$ vary greatly across $\theta \in \gamma$ for each $d \in D$ and across $d \in D$ for each $\theta \in \gamma$.

The third consideration is of course the relative cost of eliciting probabilities on different sets.

To conclude this discussion, we raise the question of the convergence of the sequential elicitation process. Observe first that for all $d \in D$

$$\Omega_0 \subset \Omega_1 \Rightarrow \underline{V}_0(d) \leq \underline{V}_1(d),\ \bar{V}_0(d) \geq \bar{V}_1(d).$$

From this it follows that

$$\infty > \bar{V}_0(\bar{d}_0) - \underline{V}_0(\underline{d}_0) \geq \bar{V}_1(\bar{d}_1) - \underline{V}_1(\underline{d}_1) \geq 0,$$

that is, the computable upper bound on the value of probability elicitation cannot increase as the domain becomes more refined.

It is of interest to seek conditions on the utilities, the action and model spaces, and the elicitation process that guarantee that this bound converges to zero as refinement proceeds. A very simple such case is that in which the sets D and Θ are finite. When the bound does converge to zero, a point must eventually come at which K exceeds the bound and hence further refinement is unprofitable. If K is sufficiently large, this point necessarily is reached before a domain rich enough to make the utilities measurable has been attained. A general examination of the circumstances under which the terminal domain does not yield measurable utilities is not attempted here.

D. THE EFFECT OF OBSERVATIONAL EVIDENCE ON THE CHOICE OF DOMAIN. Assume now that our decision maker with initial probability domain Ω_0 observes a realization y from the observation space Y. How will this event affect his decision process?

In line with the earlier discussion, assume the existence of a σ-algebra Ω under which the utility functions $U(d/\cdot)$, $d \in D$ and the likelihood function $f(y/\cdot)$ are all measurable. Then the observation of y changes the original idealized decision problem to

$$\max_{d\in D} \frac{\int_\Theta U(d/\theta) f(y/\theta)\, dP(\theta)}{\int_\Theta f(y/\theta)\, dP(\theta)}.$$

Let us now define the lower and upper marginal likelihoods

$$\underline{f}_0(y) = \textstyle\sum_{\gamma\in\Gamma_0} (\inf_{\theta\in\gamma} f(y/\theta))\cdot P_0(\gamma),\ \bar{f}_0(y) = \sum_{\gamma\in\Gamma_0} (\sup_{\theta\in\gamma} f(y/\theta))\cdot P_0(\gamma)$$

and the lower and upper posterior expected utilities

$$\underline{V}_0(d/y) = [\textstyle\sum_{\gamma\in\Gamma_0} (\inf_{\theta\in\gamma} U(d/\theta) f(y/\theta))\cdot P_0(\gamma)]/\bar{f}_0(y)$$

$$\bar{V}_0(d/y) = [\textstyle\sum_{\gamma\in\Gamma_0} (\sup_{\theta\in\gamma} U(d/\theta) f(y/\theta))\cdot P_0(\gamma)]/\underline{f}_0(y).$$

Let the bounded cardinal utilities U be normalized so that $U(d/\theta) \geq 0$, all $d \in D$, $\theta \in \Theta$. Then as long as $\underline{f}_0(y) > 0$, $\bar{V}_0(\bar{d}_0/y) - \underline{V}_0(\underline{d}_0/y)$ places a computable finite upper bound on the value of refining Ω_0, $\bar{d}_0$ and $\underline{d}_0$ being the posterior maximax and maximin solutions. (A question not addressed here is whether this is the least upper computable bound.) As before, $\bar{V}_0(\bar{d}_0/y) - \underline{V}_0(\underline{d}_0/y) < K$ is a sufficient condition for the decision maker to retain the domain Ω_0. And as before, the bound decreases in magnitude as the probability domain becomes more refined.

It should also be clear that the earlier argument for a sequential refinement process and the discussion of considerations in selecting a minimal set to split still hold. We need only add a fourth consideration to the list, that regarding the likelihood function.

Both the between-minimal set and within-minimal set spreads of likelihood values are relevant to the refinement decision. To see the former concern, consider a situation in which the likelihood function is constant within each minimal set. In this case, the likelihood function is Ω_0-measurable, Bayes rule applies and the lower and upper posterior expected utilities become

$$\underline{V}_0(d/y) = \textstyle\sum_{\gamma\in\Gamma_0} \underline{U}(d/\gamma)\cdot P_0(\gamma/y)$$

$$\bar{V}_0(d/y) = \textstyle\sum_{\gamma\in\Gamma_0} \bar{U}(d/\gamma)\cdot P_0(\gamma/y).$$

Thus the sets γ on which the likelihood value is relatively large have higher posterior than prior probability and therefore become more important in the refinement process after the observation than before.

When the likelihood function is not constant within minimal sets, the impact of the observation on the refinement process becomes more difficult to assess. One effect of within-minimal set spread is to increase $\bar{f}_0(y)$ and decrease $\underline{f}_0(y)$, thus enlarging the bound $\bar{V}_0(\bar{d}_0/y) - \underline{V}_0(\underline{d}_0/y)$ relative to the situation with no such spread. This increase in the size of the bound makes refinement of Ω_0 relatively more attractive. However, a spread of likelihood values within each γ also interacts with the spread of utility values and affects the lower and upper expected utilities in this way. Whether this effect works to tighten or loosen the bound is case specific.

E. Conclusion. The model presented here makes a start at integrating the selection of probability domains into statistical decision theory. To complement theoretical work there is a need for behavioral analysis. How rich are the probability domains that people use in practice? In what contexts do elicitation costs constitute an important barrier to the application of rich domains? When elicitation costs are important, how do people act? These questions warrant empirical investigation.

Acknowledgment. I am grateful to Morris DeGroot, George Duncan, William Eddy, and Diane Lambert for their encouragement and helpful comments. I have also benefitted from the opportunity to present this work at a session of the NBER-NSF Conference on Bayesian Inference in Econometrics and from the critique of an anonymous referee.

REFERENCES

[1] DeGroot, M. (1970). *Optimal Statistical Decisions.* McGraw-Hill.
[2] Dempster, A. (1967). Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Statist.* **38** 325–339.
[3] Dempster, A. (1968). A generalization of Bayesian inference. *J. Roy. Statist. Soc., Series B* **30** 205–232.
[4] Fine, T. (1973). *Theories of Probability.* Academic.
[5] Good, I. J. (1962). The measure of a non-measureable set. In *Logic, Methodology and Philosophy of Science.* (E. Nagel, P. Suppes and A. Tarski, eds.). Stanford Univ. Press.
[6] Keynes, J. M. (1921). *A Treatise on Probability.* MacMillan.
[7] Lindley, D. V. (1971). *Bayesian Statistics, A Review.* SIAM.
[8] Shafer, G. (1976). *A Mathematical Theory of Evidence.* Princeton.
[9] Simon, H. (1957). *Models of Man, Social and Rational.* Wiley.
[10] Zellner, A. (1977). Maximal data information prior distributions. In *New Developments in the Applications of Bayesian Methods.* (A. Aykac and C. Brumat, eds.). North-Holland.

Department of Economics
The Hebrew University of Jerusalem
Jerusalem, Israel