Review of Economic Studies (1993) 60, 487-493
0034-6527/93/00250487$02.00

Savage's Axioms Usually Imply Violation of Strict Stochastic Dominance

PETER WAKKER
University of Leiden

First version received August 1990; *final version accepted September* 1992 (*Eds.*)

Contrary to common belief, Savage's axioms do not imply strict stochastic dominance. Instead, they usually involve violation of that. Violations occur as soon as the range of the utility function is rich enough, e.g. contains an interval, and the probability measure is, loosely speaking, "constructive". An example is given where all of Savage's axioms are satisfied, but still strict statewise monotonicity is violated: An agent is willing to exchange an act for another act that with certainty yields a strictly worse outcome. Thus book can be made against the agent. Weak stochastic dominance and weak statewise monotonicity are always satisfied, as well as strict stochastic dominance and strict statewise monotonicity when restricted to acts with finitely many outcomes.

Decision making under uncertainty studies situations where an agent must make decisions in the face of uncertainty. The most famous contribution to this field is, undoubtedly, Savage (1954). He laid down, more clearly than his predecessors, the basic model for decision making under uncertainty, and formulated and axiomatized the most fruitful approach, i.e. subjective expected utility (SEU). In his model uncertainty is modelled through a *state space S*. Exactly one state is true, the others are not true, and the agent is uncertain about which state is the true one. The classical example concerns the case of a horse race, and an agent who wishes to bet on this race. For every horse there is exactly one state (i.e. element of S), describing the event that the horse in question will win the race. The agent does not know for sure which horse will win, thus is uncertain about which of the states is the true one. A second primitive in the model of Savage is the *outcome* space (or consequence space), denoted by $\mathscr{C}$ in this paper. Outcomes describe the result of the decision situation, after the agent has taken his decision and the uncertainty has been resolved. In the example of the horse race, we assume that outcomes are real numbers, designating amounts of money. So here it is assumed that the only relevant aspect of the outcome of the decision situation is the net gain (or loss) that results for the agent.

The agent must choose between a set of available decision alternatives, called *acts* in this context. In the example of the horse race the available acts will be stakes, or combinations thereof. Formally an act is described as a function from S to $\mathscr{C}$. So an act is described by the outcome that it will yield for every state. If an act is a combination of stakes on the horse race, then it is described by the net gain that will result for every state. Note that this is a consequentialistic approach, where only final outcomes matter. As the agent is uncertain about which state is the true one, he is uncertain which outcome will result from a chosen act. The preference relation of the agent over the acts is described by a binary relation $\succcurlyeq$ on the set of acts. Savage's SEU model prescribes that the agent choose his utility function $U: \mathscr{C} \rightarrow \mathbb{R}$, and his ("subjective") probability measure P on 2^S, describing for every (event =) subset of the state space the likelihood, according to the agent's beliefs, that the event contains the true state. Then the agent will prefer acts according to their SEU value. That is, $f \succcurlyeq g \Leftrightarrow \int_S (U \circ f) dP \geqq \int_S (U \circ g) dP$.

Observable primitives in this setup are, besides the acts described as outcomes contingent on the states, the preferences of the agent over the acts. Probabilities and utilities are not primitives. They are theoretical constructs. Thus, without further justification, the claim that SEU will or should be maximized is meaningless. The great achievement of Savage was to present a list of seven axioms, directly in terms of preferences, that were subsequently shown to imply SEU maximization. Thus the status of SEU changed from being ad hoc to being empirically meaningful. Savage was the first to complete such a result. An earlier attempt by Ramsey (1931) was incomplete, the result of de Finetti (1937), underestimated in the economic literature, had the drawback that utility was assumed to be known beforehand, and the result of von Neumann and Morgenstern (1944) had the drawback that probability was assumed to be known beforehand.

The most important postulates of Savage were $P2$ and $P4$. $P2$, the "sure-thing" principle, requires that the preference between two acts conditional on an event A be independent of what would have happened if A would not have obtained. $P4$ requires for instance that the preference of receiving \$1 contingent on event A over receiving \$1 contingent on event B should imply the preference of receiving \$2 contingent on event A over receiving \$2 contingent on event B. As with all axiomatizations, Savage's does not only invoke intuitive postulates, but also technical ones, that for the purpose of this paper need not be spelt out.

Savage's work gave a foundation to Bayesian statistics, and to the modelling of uncertainty by SEU in economics. It also showed how to construct critical empirical tests of SEU. These have led to many criticisms of SEU, and alternative models, in the recent literature on decision theory.[1] The implications of Savage's technical postulates have not been well-understood in the literature. Authors, when using SEU models, usually refer to Savage for a justification, despite the fact that their models do not satisfy Savage's technical postulates. This paper aims to contribute to a further understanding of the technical complications in Savage's approach, in particular concerning the finite additivity of his probability. The paper will show that strict stochastic dominance, a condition generally accepted in economic theory, is usually excluded by Savage's technical postulates. For other comments on Savage's approach, see Wakker (1993).

It follows directly from SEU maximization that $\succcurlyeq$ satisfies the *reduction principle*, i.e. if two acts generate the same probability distribution over $\mathscr{C}$ then they are equivalent. Hence one may identify an act with the probability distribution it generates over the outcomes. Among the hallmarks of expected utility theory are the results on risk aversion of Pratt (1964) and Arrow (1965), and on the impact of increases in risk for the demands of risky assets in Rothschild and Stiglitz (1970, 1971), leading to notions of higher-order stochastic dominance. The most elementary, and universally accepted, form of stochastic dominance is first-order stochastic dominance (or stochastic dominance for short). When models were proposed in the literature that turned out to violate stochastic dominance, these models were generally criticized as being irrational. See, for instance, the model

1. See for instance Machina (1987) and Fishburn (1988).

proposed in Handa (1977), and subsequently criticized in Fishburn (1978). This led Quiggin (1982) to introduce what is now the most popular deviation from expected utility: rank-dependent utility (=anticipated utility). Also Kahneman and Tversky (1979) had to complicate their famous prospect theory so as to avoid violation of stochastic dominance. This has recently led to a new version of prospect theory in order to overcome this complication; see Tversky and Kahneman (1992) and Wakker and Tversky (1991). Machina (1982) generalized Savage's expected utility to almost complete generality, but emphatically did preserve the implication of stochastic dominance. The condition is also central in Machina and Schmeidler (1990). So far it has remained unnoticed in the literature that the technical conditions of Savage, under a minimal richness condition and a constructivity condition, actually necessitate a violation of stochastic dominance! Let me emphasize that this is merely an implication of Savage's technical postulates, and not of his, more important, intuitive conditions. At the end of the paper it will be explained how the problem can be amended.

We now turn to definitions. $F_{U\circ f}$ denotes the distribution function of $U\circ f$, i.e. $F_{U\circ f}: \tau \mapsto P(U\circ f \leqq \tau)$. We say that $F_{U\circ f}$ *weakly stochastically dominates* $F_{U\circ g}$ if $F_{U\circ f}(\tau) \leqq F_{U\circ g}(\tau)$ for all τ; *strict stochastic dominance* holds if the two distribution functions are not identical. Savage's axioms do straightforwardly imply (*monotonicity with respect to*) *weak stochastic dominance*, i.e. $f \succcurlyeq g$ whenever $F_{U\circ f}$ weakly stochastically dominates $F_{U\circ g}$. The definition of stochastic dominance, however, entails the following stronger condition: $\succcurlyeq$ satisfies *strict* (*monotonicity with respect to*) *stochastic dominance* if $f \succ g$ whenever $F_{U\circ f}$ strictly stochastically dominates $F_{U\circ g}$.

A (*finitely additive*) probability measure P satisfies $0 \leqq P(A) \leqq 1$ for all A, $P(A \cup B) = P(A) + P(B)$ for all disjoint A, B (finite additivity), and $P(S) = 1$. By induction then

$$P(\cup_{i=1}^{n} A_j) = \sum_{i=1}^{n} P(A_j)$$

for any finite disjoint sequence $A_1, \ldots, A_n$. However, in general the equality

$$P(\cup_{i=1}^{\infty} A_j) = \sum_{i=1}^{\infty} P(A_j)$$

does not have to hold true for all infinite disjoint sequences $A_1, A_2, \ldots$ If the equation does hold true, then P is called *countably additive.* The difference between finite additivity and countable additivity is central in this paper. It is well-known that SEU maximization implies strict stochastic dominance for countably additive probabilities P on S. As we shall see, this is not the case for finitely additive P.

To avoid measure-theoretic complications, Savage assumed that the domain of P is the power set 2^S. He pointed out that any σ-algebra of subsets of S could be taken as well. (Savage's analysis does not hold for general algebras; for comments see Wakker (1981).) On a general σ-algebra, P could be countably additive indeed, in which case strict stochastic dominance would be satisfied. The existence of a countably additive P is more complicated for the special case where the σ-algebra is 2^S. Under the generally accepted set-theoretic axioms it is an undecidable problem whether such a P can exist; the answer to the question depends on the specific set-theoretic axioms that one adds to the usual ones. This is not the right place (or author) to enter a detailed discussion of these; details can be found in Jech (1978, in particular Chapter 5). Suffice it to say that the question is open. A simple example of an axiom that excludes the existence of a countably additive atomless P, is the "constructibility axiom"; see Scott (1961).[2] This

2. The constructibility axiom is often denoted as "$V = L$". See also Jech (1978, beginning of Section 31). Under the axiom of determinacy, to the contrary, such a P can be obtained. See Mycielski and Swierczkowski (1964), or Jech (1978, Theorem 102a).

means that, if such a P exists, then loosely speaking, one will never "find" it. For finitely additive probability measures, strict stochastic dominance is more problematic than has often been thought. This paper shows that strict stochastic dominance cannot hold true in Savage's model as soon as the following richness assumption is satisfied, in addition to the needed set-theoretic axioms.

Assumption 1. The range of U contains a strictly decreasing sequence $(U(\alpha_j))_{j=1}^{\infty}$ as well as its limit $U(\alpha)$.

Strictly speaking, Assumption 1 is not a proper behavioural condition, because it is stated in terms of utility, which is not a directly observable primitive. The following assumption gives a proper behavioural reformulation of Assumption 1. The verification of the equivalence of Assumptions 1 and 2 is by substitution of SEU.

Assumption 2. There exist outcomes α and $(\alpha_j)_{j=1}^{\infty}$ such that $\alpha_j \succ \alpha_{j+1} \succ \alpha$ for all j, and there exists an event A such that for each α_j there exists α_{j+k} such that α_j is strictly preferred to the act which assigns α to A and α_j to A^c (to ensure that A is "non-null"), and the latter act is strictly preferred to α_{j+k}.

Theorem A1 in the appendix will show that Savage's postulates necessitate a violation of strict stochastic dominance as soon as Assumption 1, or 2, is satisfied, and some set-theoretic axioms. Corollary 3 below gives a more accessible version of the result. Decision models should be able to deal with the case considered in the Corollary. Hence strict stochastic dominance is problematic in Savage (1954).

Corollary 3. *If the range of the utility function contains a non-degenerate interval in Savage's model, then, under the generally accepted set-theoretic axioms, it cannot be decided if strict stochastic dominance can ever be satisfied; if the set-theoretic constructibility axiom is accepted, then strict stochastic dominance is necessarily violated.*

While Machina (1982) showed under countable additivity that stochastic dominance results still hold for very general non-expected utility models, this paper has shown that under finite additivity they do not even hold under expected utility.

One way out of the problem is to impose strict stochastic dominance only on the *simple* acts, i.e. the acts that have a finite range. Another way out is to re-define strict stochastic dominance: one distribution strictly stochastically dominates another only if it weakly stochastically dominates the other and further there exists a positive ε and a positive δ such that on an interval of length at least δ the dominated distribution function is at least ε larger than the dominating one. Under countable additivity this definition coincides with the traditional one because of right continuity of distribution functions, under finite additivity it avoids the violation of strict stochastic dominance in Savage's model. So this may be the proper general definition of strict stochastic dominance.

A similar misunderstanding about Savage's model is the following: Contrary to what has often been thought, strict statewise monotonicity of the form $[f(s) \succ g(s)$ for all $s \in S \Rightarrow f \succ g]$ need not hold. The example below illustrates. Again, only weak statewise monotonicity holds, and strict statewise monotonicity still holds when restricted to simple acts, or to the case where P is countably additive. The example also shows that Savage's state space may be countable, contrary to what has sometimes been thought.

Example 4. Let $S=[0,1]\cap\mathbb{Q}$ be a countable state space ($\mathbb{Q}$ is the set of rationals). $\mathscr{C}=[0,1]$, and utility is linear. Suppose for each interval $[\mu,\nu]$ within $[0,1]$ (so also with μ, ν irrational) $P([\mu,\nu]\cap S)=\nu-\mu$. P can be extended to all subsets of S, with all axioms of Savage satisfied (e.g. by the Theorem of Hahn-Banach from functional analysis, see Dunford and Schwartz (1958)). Obviously, P assigns value 0 to each finite event. The states can be numbered as $(s_j)_{j=1}^{\infty}$. Let $f: s_j \to 1/j$, $g: s_j \to 1/(j+1)$. Both acts have a distribution function that is 0 on $]-\infty,0]$, 1 on $]0,\infty]$, i.e. they have the same distribution function. Hence they are equivalent. But $f(s)>g(s)$ for all s.

In the above example the agent, while conforming to Savage's principles, is willing to exchange the act f for an act g, whereas g results with certainty in a strictly worse outcome. In that sense the agent is vulnerable to immediate book making, that does not invoke combinations of decisions or repeated exchanges of acts as in de Finetti (1937) or Green (1987).

Let me emphasize that the findings of this paper are not meant to cast doubt on Savage's axioms. Rather they are meant to show that one should be careful in extending findings from the usual countably additive probabilities to finitely additivity probabilities. A way out of the problem as pointed out in the above example is to impose strict statewise monotonicity only on simple acts. Another way out is by saying that act f strictly statewise dominates act g only if it satisfies one further restriction: There should exist outcomes $\alpha>\beta$ and an event with strictly positive probability, such that on that event $f(s)\geq\alpha>\beta\geq g(s)$. In the usual countably additive contexts it can be seen that this condition is not additionally restrictive, whereas in finitely additive contexts such as Savage's model the condition is also satisfied, so that paradoxes as above are avoided.

The axiomatizations of SEU in Pratt, Raiffa and Schlaifer (1964), as well as in Anscombe and Aumann (1963), assumed, more or less similar to von Neumann and Morgenstern (1944), that (some) probabilities were known. The axiomatizations of SEU in Wakker (1989, Theorem IV.2.7) and Gul (1989) avoided the technical complications of Savage, and led to setups where $u(\mathscr{C})$ is indeed an interval. Like Savage, they assume neither probabilities nor utilities given in advance.

Quiggin (1989) studied stochastic dominance and pointwise monotonicity for regret theory. He found violations of stochastic dominance, even in the weak sense. Violations of weak statewise monotonicity (called "dominance" by Quiggin) do not occur under regret theory.

Conclusion 5. In contexts where countable additivity is not ensured, it seems more natural to impose only weak statewise monotonicity/stochastic dominance, and to restrict strict statewise monotonicity/stochastic dominance, for example to the simple acts/probability distributions.

APPENDIX

Theorem A1. *If the seven postulates of Savage (1954) hold, as well as Assumption 1 (or, equivalently, 2), then under the usual set-theoretic axioms, it cannot be decided if at all strict stochastic dominance can be satisfied. If the constructibility axiom is added to the set-theoretic axioms, then strict stochastic dominance is necessarily violated.*

Proof. We assume below that the probability measure P is strictly finitely additive, and then derive a violation of strict stochastic dominance. As the constructibility axiom implies that P must be strictly finitely additive, and under the usual set-theoretic axioms it cannot be decided if P is necessarily strictly finitely additive, this proves the theorem.

Note that the reasoning below applies to *any* strictly finitely additive probability measure P, also if the domain would be different from 2^S. There must exist a countable number of disjoint events $(A_j)_{j=1}^{\infty}$ such that $\sum_{j=1}^{\infty} P(A_j) < P(\cup_{j=1}^{\infty} A_j)$. Note that $>$ instead of $<$ can never occur, because $\sum_{j=1}^{n} P(A_j) = P(\cup_{j=1}^{n} A_j) \leqq P(\cup_{j=1}^{\infty} A_j)$ for each n. Let f be the act that assigns α_j to each A_j, and α to the remaining states, with α_j and α as in Assumption 1 or 2. Savage's axioms imply that the probability measure P on 2^S is *convex-ranged*, i.e. for each $A \subset S$ and $0 \leqq \beta \leqq P(A)$ there exists $B \subset A$ with $P(B) = \beta$. This implies that P is atomless. By convex-rangedness, we can take an event B such that $P(B) = \sum_{j=1}^{\infty} P(A_j)$, we can take an event $B_1 \subset B$ such that $P(B_1) = P(A_1)$, and by induction we can take for every j an event $B_j \subset B \backslash (\cup_{i=1}^{j-1} B_i)$ such that $P(B_j) = P(A_j)$. Obviously, $P(B) \geqq P(\cup_{j=1}^{\infty} B_j) \geqq \sum_{j=1}^{\infty} P(B_j) = \sum_{j=1}^{\infty} P(A_j) = P(B)$, so that all inequalities are in fact identities. In particular, $P(B \backslash \cup_{j=1}^{\infty} B_j) = 0$. We may, and will, assume that $B \backslash \cup_{j=1}^{\infty} B_j = \varnothing$, which can be obtained by replacing B_1 by $B_1 \cup (B \backslash \cup_{j=1}^{\infty} B_j)$. Let g be the act that assigns α_j to each B_j, and α to the remaining states. The distribution functions $F_{U \circ f}$ and $F_{U \circ g}$ are identical except at the point $U(\alpha)$: We have $F_{U \circ f}(U(\alpha)) = 1 - P(\cup_{j=1}^{\infty} A_j) < 1 - \sum_{j=1}^{\infty} P(A_j) = 1 - \sum_{j=1}^{\infty} P(B_j) = 1 - P(\cup_{j=1}^{\infty} B_j) = F_{U \circ g}(U(\alpha))$. So f strictly stochastically dominates g. Nevertheless $f \sim g$, because both acts have the same SEU value, e.g. because their distribution functions dominate/are dominated by the same "simple" distribution functions. (*Simple* means that probability 1 is assigned to some finite outcome set.) Note that f generates a not-countably additive probability distribution over $\mathscr{C}$ and a distribution function $F_{U \circ f}$ that is not right-continuous at α, whereas g generates a countably additive probability distribution over $\mathscr{C}$ and a distribution function $F_{U \circ g}$ that is right-continuous. ||

Acknowledgement. The research has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences, and a fellowship of the Netherlands Organization for Scientific Research. The author is grateful for the helpful comments of three anonymous referees.

REFERENCES

ANSCOMBE, F. J. and AUMANN, R. J. (1963), "A Definition of Subjective Probability", *Annals of Mathematical Statistics*, **34**, 199-205.

ARROW, K. J. (1965) *Aspects of the Theory of Risk-Bearing* (Helsinki: Academic Bookstore). (Elaborated as Arrow, K. J. (1971) *Essays in the Theory of Risk-Bearing.* (Amsterdam: North-Holland).

DE FINETTI, B. (1937), "La Prévision: Ses Lois Logiques, ses Sources Subjectives", *Annales de l'Institut Henri Poincaré*, **7**, 1-68. (Translated into English by H. E. Kyburg, "Foresight: Its Logical Laws, its Subjective Sources", in H. E. Kyburg and H. E. Smokler (Eds.), *Studies in Subjective Probability*, (First edition 1964, Wiley, New York, 53-118; Second edition 1980, Krieger, New York).

DUNFORD, N. and SCHWARTZ J. T. (1958) *Linear Operators, Part I* (New York: Interscience Publishers).

FISHBURN, P. C. (1978), "On Handa's 'New Theory of Cardinal Utility' and the Maximization of Expected Return", *Journal of Political Economy*, **86**, 321-324.

FISHBURN, P. C. (1988) *Nonlinear Preference and Utility Theory* (Baltimore: Johns Hopkins University Press).

GREEN, J. (1987), "'Making Book against Oneself', The Independence Axiom, and Nonlinear Utility Theory", *Quarterly Journal of Economics*, **102**, 785-796.

GUL, F. (1989), "Savage's Theorem with a Finite Number of States" (mimeo, Stanford University, Stanford).

HANDA, J. (1977), "Risk, Probabilities, and a New Theory of Cardinal Utility", *Journal of Political Economy*, **85**, 97-122.

JECH, T. (1978) *Set Theory* (New York: Academic Press).

KAHNEMAN, D. and TVERSKY, A. (1979), "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, **47**, 263-291.

MACHINA, M. J. (1982), "Expected Utility Analysis without the Independence Axiom", *Econometrica*, **50**, 277-323.

MACHINA, M. J. (1987), "Choice under Uncertainty: Problems Solved and Unsolved, *Journal of Economic Perspectives*, **1**, 121-154.

MACHINA, M. J. and SCHMEIDLER, D. (1990), "A More Robust Definition of Subjective Probability", *Econometrica* (forthcoming).

MYCIELSKI, J. and SWIERCZKOWSKI, S. (1964), "On the Lebesgue Measurability and the Axiom of Determinateness", *Fundamenta Mathematicae*, **54**, 67-71.

PRATT, J. W. (1964), "Risk Aversion in the Small and in the Large", *Econometrica*, **32**, 122-136.

PRATT, J. W., RAIFFA, H. and SCHLAIFER, R. (1964), "The Foundations of Decision under Uncertainty: An Elementary Exposition", *Journal of American Statistical Association*, **59**, 353-375.

QUIGGIN, J. (1982), "A Theory of Anticipated Utility", *Journal of Economic Behaviour and Organization*, **3**, 323-343.

QUIGGIN, J. (1989), "Stochastic Dominance in Regret Theory", *Review of Economic Studies*, **57**, 503-511.

RAMSEY, F. P. (1931), "Truth and Probability". In his *Foundations of Mathematics and other Logical Essays*, 156-198 (London: Routledge and Kegan Paul). (Reprinted in H. E. Kyburg and H. E. Smokler (Eds.), *Studies in Subjective Probability*, 61-92 (New York: Wiley) (1964).)

ROTHSCHILD, M. and STIGLITZ, J. (1970), "Increasing Risk: I. A Definition", *Journal of Economic Theory*, **2**, 225-243.

ROTHSCHILD, M. and STIGLITZ, J. (1971), "Increasing Risk: II. Its Economic Consequences", *Journal of Economic Theory*, **3**, 66–84.
SAVAGE, L. J. (1954) *The Foundations of Statistics* (New York: Wiley) (Second edition 1972, Dover, New York.)
SCOTT, D. (1961), "Measurable Cardinals and Constructible Sets", *Bulletin de l'Académie Polonaise des Sciences*, **9**, 521–524.
TVERSKY, A. and KAHNEMAN, D. (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty", *Journal of Risk and Uncertainty* (forthcoming).
VON NEUMANN, J. and MORGENSTERN, O. (1944, 1947, 1953) *Theory of Games and Economic Behavior* (Princeton NJ: Princeton University Press).
WAKKER, P. P. (1981), "Agreeing Probability Measures for Comparative Probability Structures", *The Annals of Statistics*, **9**, 658–662.
WAKKER, P. P. (1989) *Additive Representations of Preferences, A New Foundation of Decision Analysis* (Dordrecht: Kluwer Academic Publishers).
WAKKER, P. P. (1993), "Clarification of some Mathematical Misunderstandings about Savage's Foundations of Statistics, 1954", *Mathematical Social Sciences*, **25**, (forthcoming).
WAKKER, P. P. and TVERSKY, A. (1991), "An Axiomatization of Cumulative Prospect Theory" (mimeo, University of Nijmegen, NICI, Nijmegen, The Netherlands).