Review of Economic Studies (1997) **64**, 125–146 0034-6527/97/00060125$02.00

Countably Additive Subjective Probabilities

MAXWELL B. STINCHCOMBE
University of Texas, Austin

First version received November 1993; *final version accepted August* 1996 (*Eds.*)

The subjective probabilities implied by Savage's (1954, 1972) Postulates are finitely but not countably additive. The failure of countable additivity leads to two known classes of dominance paradoxes, money pumps and indifference between an act and one that pointwise dominates it. There is a common resolution to these classes of paradoxes and to any others that might arise from failures of countably additivity. It consists of reinterpreting finitely additive probabilities as the "traces" of countably additive probabilities on larger state spaces. The new and larger state spaces preserve the essential decision-theoretic structures of the original spaces.

1. INTRODUCTION

Savage's (1954, 1972) framework for modelling choice under uncertainty provides a theory of subjective probability, and has been called the "crowning glory of choice theory." (Kreps (1988)). Any adequate education in modern economics must include his Subjective Expected Utility (SEU), this despite the paradoxes Savage's framework is known to contain. The paradoxes, two kinds of money pumps or Dutch book, and indifference between an act and one that pointwise dominates it do not depend on violations of the Sure Thing Principle. Rather, they arise because Savage chose to work in a framework that implies that the subjective probabilities fail to be countably additive.[1] Reinterpreting finitely additive probabilities as the "traces" of countably additive probabilities on larger state spaces resolves these paradoxes by showing that they ignore sets of positive probability. The resolution of the paradoxes can only be as strong as the match between the original finitely additive decision framework and its countably additive reinterpretation.

1.1. *The match*

Savage's decision framework consists of a state space S and a set of consequences $\mathbf{C}$. A *gamble*, or *simple act*, is any mapping $g: S \to \mathbf{C}$, the range of which is finite, and an *act* is any mapping $a: S \to \mathbf{C}$. Preference relations, $\preceq$, are defined on the set of acts. If the preference relation $\preceq$ satisfies six Postulates (in Appendix A for ease of reference), then it has an SEU representation for simple acts (gambles), that is, there exists a subjective probability P defined on S and a bounded, real-valued function U defined on $\mathbf{C}$ such that $g_1 \preceq g_2$ if and only if $\int_S U(g_1(s))dP(s) \leq \int_S U(g_2(s))dP(s)$. If $\preceq$ satisfies an additional seventh Postulate, it has an SEU representation for acts.

Suppressing some details, a reinterpretation of a Savage decision framework is an embedding of the state space S in a larger space $\hat{S}$, extensions of the acts and gambles

1. Adams (1962) and Seidenfeld and Schervish (1983) give examples of money pumps and indifference between an act and a pointwise dominant act, Wakker (1993) gives an example of the latter.

from $S \subset \hat{S}$ to all of $\hat{S}$, $g \mapsto \hat{g}$ and $a \mapsto \hat{a}$, and an extension of P to $\hat{P}$ defined on the appropriate collection of subsets of $\hat{S}$. The match between the original framework and its reinterpretation is quite close.

The embeddings considered here produce countably additive $\hat{P}$ that are determined by the fact that they are extensions of P—that for all $E \subset S$, $\hat{P}(\hat{E}) = P(E)$. The new acts, when applied to $\hat{P}$, produce the same distributions on the set of consequences. Again suppressing some details, the image distribution of $\hat{P}$ induced by $\hat{a}$ is essentially indistinguishable[2] from the image distribution of P induced by a. For probabilistically sophisticated preferences (Machina and Schmeidler (1992)), that is, preferences over A that depend only on the distributions induced on consequences, the reinterpretations provided here provide equivalent frameworks for choice under uncertainty. Thus, Machina (1982) locally linear or Quiggin (1993) rank-dependent preferences over finitely additive probabilities can be reinterpreted as preferences over countably additive probabilities. In particular, one aspect of distributions being indistinguishable is that their integrals are the same, $\int_S U(a(s))dP(s) = \int_{\hat{S}} U(\hat{a}(\hat{s}))d\hat{P}(\hat{s})$.

1.2. *The mismatch*

Savage makes two assumptions on the framework, which, taken together, force the subjective probability P to be finitely but not countably additive. First, the state space S is infinite and the probability P is non-atomic (in the finitely additive sense that it allows for partitions of the state space into sets of arbitrarily small probability). This is essential to the theory because it is necessary to partition S arbitrarily finely in order to obtain arbitrary probabilities. Second, all acts or gambles are evaluated, not just a subset of acts or gambles measurable with respect to some σ-field of subsets of S smaller than the set of all subsets of S. A non-atomic P defined on all subsets of an infinite S must fail countable additivity.[3]

As Savage indicates (1954, 1972, Section 3.4), the use of the set of all subsets is not necessary for his development. However, it does have an important implication for the interpretation of what he wrote. The restriction that gambles (or acts) be measurable with respect to some σ-field of subsets of S smaller than the set of all subsets of S is clearly a restriction on the set of gambles (or acts) being considered. Despite this, there is no state at which the restriction is binding—changing any gamble (or act) at a single state results in another measurable gamble (or act).[4] Put another way, the restriction to measurable gambles (or acts) gives pointwise absolute freedom even though overall choices are constrained. By allowing all gambles (or acts), Savage allows for both pointwise absolute freedom and overall absolute freedom.

The resolution to the paradoxes that is proposed here loses this overall absolute freedom. The new, larger state spaces have non-atomic, countably additive subjective

2. This means exactly equal for simple acts, and weak*-equivalent for general acts.

3. In their discussion of the appropriate space of strategies for non-atomic games, Aumann and Shapley (1978) note that Sierpinski (1956, 1st ed. 1934) showed that if the continuum hypothesis is accepted, then a measure P must be purely finitely additive if it is non-atomic on the set of all subsets of an infinite set. Savage (1954, 1972) notes Ulam's (1930) proof of the same result for the unit interval. Wakker (1993) notes that the same conclusion holds if the stronger Axiom of Constructibility is accepted. de Finetti (1972, 1974, 1975) uses the fact that countably additive measures necessarily entail nonmeasurable sets as a major argument against countable additivity. Both the continuum hypothesis and the Axiom of Constructibility are known to be independent of the usual axioms of set theory.

4. This statement assumes a regularity condition—the σ-field contains $\{s\}$ for all $s \in S$. Passing to equivalence classes can be done so as to make this regularity condition un-needed.

probabilities on them. Therefore, the new probabilities must be defined on σ-fields that are smaller than the set of all subsets of the new spaces. A cost to working with the reinterpretations is measurability requirements. In the new state spaces, there are (non-measurable) sets of states to which the agent cannot assign a probability.[5]

A second mismatch between decision frameworks with finitely additive subjective probabilities and their countably additive reinterpretations is that the new state spaces may have peculiar mathematical properties. The leading example arises when the state space are required to be minimal. In this case, the new space is a large compactification of the original state space. Working with non-minimal reinterpretations avoids this problem.

1.3. *Summary and outline*

This paper provides a common explanation of the money pump and dominance paradoxes, and of any other paradox(es) that may arise in the future out of the failure of countable additivity. The explanation involves reinterpreting finitely additive subjective probabilities as the "traces" of countably additive subjective probabilities on larger spaces. The original state space S is regarded as being embedded in the new state space $\hat{S}$. Any paradox arising from a failure of countable additivity in S can then be re-interpreted as an ill-posed problem in $\hat{S}$. This resolution of the paradoxes is only as strong as the match between the original finitely additive decision framework and its reinterpretation. The reinterpretation reproduce an essentially equivalent version of the decision framework. The strength of this equivalence suggests (and this is discussed further in Section 8) that the state spaces $\hat{S}$ are "truer" versions of the state space. The cost of the reinterpretations is that a measurable structure is required, and $\hat{S}$ may have some peculiar properties.

The next section contains a short review of finitely additive probabilities and their differences from countably additive probabilities. The following section gives the paradoxes and intuitive versions of the ways in which the reinterpretations resolve them. Section 4 defines Bayesian decision frameworks and their reinterpretations, Section 5 gives the properties of reinterpretations. Following this, Section 6 examines how the reinterpretations resolve the paradoxes. The penultimate section, Section 7, discusses the issue of finding a minimal reinterpretation, and shows that minimal reinterpretations exist. Finally, Section 8 concludes with a methodological discussion, interpreting the new points in the state spaces, and pointing out the parallels with discussions from social choice theory, stochastic process theory, statistical decision theory, and game theory. Proofs are in Appendix C.

2. THE PARADOXES OF FINITE ADDITIVITY

This section reviews how finitely additive probabilities arise, and includes a summary of the relevant implications of the failure of countable additivity.

2.1. *Finitely additive probabilities*

A canonical probability that fails countable additivity can be had by trying to find a "uniform" distribution over the integers. For each $n \in \mathbb{N}$, let λ_n be the uniform distribution on $\{1, 2, \ldots, n\}$. For any $E \subset \mathbb{N}$, $(\lambda_n(E))_{n \in \mathbb{N}} \subset [0, 1]$ has at least one accumulation point, and there is (Alaoglu's Theorem) a finitely additive λ on the set of all subsets of the integers that consistently picks from these accumulation points, essentially by taking convergent

5. Skyrms (1995) and also (1993) discusses these and related issues.

subsequences.[6] Thus, $\lambda\{\text{evens}\}=1/2$ because $\lim_n \lambda_n\{\text{evens}\}=1/2$, $\lambda\{7m: m\in\mathbb{N}\}=1/7$, $\lambda\{2^m: m\in\mathbb{N}\}=0$, and there are sets E for which $(\lambda_n(E))_{n\in\mathbb{N}}$ has many accumulation points.

2.2. *Relevant properties of* λ

1. λ is finitely additive—for any disjoint A and B, $\lambda_n(A\cup B)=\lambda_n(A)+\lambda_n(B)$, and this property is preserved at accumulation points.
2. λ fails to be countably additive—let $E_n=\{m\in\mathbb{N}: m\leqq n\}$, $E_n\uparrow\mathbb{N}$, yet $\lambda(E_n)\equiv 0$, so that $\lim_n \lambda(E_n)<\lambda(\lim_n E_n)$.
3. Any finite E satisfies $\lambda(E)=0$.
4. λ is non-atomic—for any $\varepsilon>0$, it is possible to partition $\mathbb{N}$ into finitely many sets E_i with $\lambda(E_i)<\varepsilon$.
5. Any bounded function g on $\mathbb{N}$ is λ-integrable, and the integral can be defined by

$$\int_{\mathbb{N}} g(n)d\lambda(n)=\lim_{m\uparrow\infty}\sum_{i=-m2^m}^{+m2^m}\frac{i}{2^m}\lambda\left\{g\in\left[\frac{i}{2^m},\frac{i+1}{2^m}\right)\right\}. \tag{1}$$

2.3. *Dominated convergence fails*

The failure of countable additivity is equivalent to the failure of Lebesgue's dominated convergence theorem. To see why this is true, let g_n be the indicator of the set E_n in the second point above. Then $g_n(m)\uparrow 1$ for each $m\in\mathbb{N}$, yet $\lim_n \int g_n d\lambda=0<\int \lim_n g_n d\lambda=1$. This failure is at the heart of Adams' (1962) money pump, and is one way to understand Seidenfeld and Schervish's (1983) money pump.

2.4. *The integral of a strictly positive function may be zero*

The failure of countable additivity also allows for two functions to satisfy $f(m)>g(m)$ for each $m\in\mathbb{N}$, yet to have the same integral. In particular, suppose that for all $m\in\mathbb{N}$, $f(m)>g(m)\geqq 0$, and that $\lim_{m\to\infty}f(m)=0$. Then

$$\int_{\mathbb{N}} f(m)d\lambda(m)=\int_{\mathbb{N}} g(m)d\lambda(m)=0 \tag{2}$$

because $\lambda\{f,g\in[0,1/2^m)\}\equiv 1$. This property is at the heart of the dominance paradoxes.

2.5. *Conglomerability fails*

A more subtle property equivalent to countable additivity is *conglomerability* (de Finetti (1972), Dubins (1975), Seidenfeld, Schervish and Kadane (1984), Armstrong (1990)). A probability P *fails* conglomerability in a countable partition $\pi=\{E_1, E_2,\ldots\}$ of a state space S, if there is some event E, and constants $k_1\leqq k_2$ such that $k_1\leqq P(E|E_i)\leqq k_2$ for each $E_i\in\pi$, yet $P(E)<k_1$ or $P(E)>k_2$. Failing conglomerability means that there is an event E, and a partition π with the property that, conditional on each and every event in π, the posterior probability of E is above (or below) the prior probability of E. It is known

6. More formally, note that probabilities are determined as points in $X:=\times_{E\subset\mathbb{N}}[0,1]$. With the product topology, this is a compact space. The infinite set $\{(\lambda_n(E))_{E\subset\mathbb{N}}: n\in\mathbb{N}\}\subset X$ must therefore have accumulation points. The work in proving Alaoglu's theorem (e.g. Royden (1968, Theorem 10.17, p. 202)) is in showing that any such accumulation point is a finitely additive probability.

that a finitely additive probability fails conglomerability in some partition if and only if it fails to be countably additive.

The decision-theoretic implications of a failure of conglomerability can be striking. If conglomerability fails, there is an event E, a countably infinite partition π, and a bet conditional on E that an agent will pay to take, even though, conditional on each and every $E_i \in \pi$, they are willing to pay to get out of. This property is at the heart of Seidenfeld and Schervish's (1983) money pump.

2.6. *Bayesian statistical interpretations fail*

The following is an implication of the failure of conglomerability taken from Heath and Sudderth (1989). Suppose that the parameter space is $\Theta = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ in a statistical model, and that if $\theta \in \Theta$ is true, the observation X will be made according to a distribution P_θ. Suppose that $P_\theta(X = \theta - 1) = P_\theta(X = \theta + 1) = 1/2$. This means that if θ is an even number, the observation will be an odd number and *vice versa*.

The statistician's prior distribution over Θ is $P = 1/2Q + 1/2\mu$ where Q is a countably additive probability with $Q(\theta) > 0$ for each $\theta \in \Theta$, and $Q\{\text{evens}\} = 1/2$. The probability μ is any one of the purely finitely additive accumulation points (in $\times_{E \subset \Theta}[0, 1]$) of the sequence

$$\mu_n = \text{Unif}\,\{-2n, -2(n-1), \ldots, -2, 0, +2, \ldots, 2(n-1), 2n\}, \tag{3}$$

so that $\mu\{\text{evens}\} = 1$.

If the statistician sees the event $\{\text{odds}\}$, the posterior distribution is $1/3Q + 2/3\mu$. On the other hand, if the event $\{m\}$ is observed, the posterior distribution is $Q(\cdot \mid \{m-1, m+1\})$ because $Q(\{m-1, m+1\}) > 0$ while $\mu(\{m-1, m+1\}) = 0$. The posterior ignores the μ part of the prior distribution for every realization $m \in \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$. This shows that "drawing" a θ according to P does not have some of the intuitive properties of random draws. Here, the half of the mass described by μ seems to be lost. One interpretation is that the half of the mass described by μ occurs, but as new points in some larger $\hat{\Theta}$, a parameter space that contains ideal points that are to the "right" and to the "left" of the set Θ.

2.7. *Compactifications*

Compactifications add ideal points to sets. Note that the c.d.f.'s G_n of λ_n or the c.d.f.'s H_n of μ_n are not tight in the sense of the weak convergence of probability measures (e.g. Billingsley (1979, Section 25)). In particular, for each number $r \in \mathbb{R}$, $G_n(r)$, $H_n(r) \downarrow 0$ as the mass in λ_n "moves to the right" and the mass in μ_n moves away from 0 in $\mathbb{R}$. One response to this might be to interpret the limit, λ, as putting all of its mass on the set $\{+\infty\}$ where ∞ is an ideal point added to the "right" of $\mathbb{R}$. In a similar fashion, μ might be interpreted as putting half its mass on $\{-\infty\}$ and half on $\{+\infty\}$. In other words, regarding $\mathbb{R} = (-\infty, +\infty)$ as a dense subset of the compact space $[-\infty, +\infty]$, the limits of the λ_n and the μ_n are distributions on the new, larger state space.

This simple compactification of the space $\mathbb{R}$ will not work for our purposes—note that in reducing λ to a point mass, it loses the property of being non-atomic. One requirement of the state spaces used in the reinterpretations is that they are sufficiently rich that the limit of the λ_n is again a non-atomic distribution.

3. THE PARADOXES

This section begins with an example due, in slightly different forms, to Adams (1962), Seidenfeld and Schervish (1983), and Wakker (1993). An agent may have SEU preferences with a continuous expected utility function, yet be indifferent between an act and a pointwise dominant act. Following this are the two rather different money pump constructions of Adams (1962) and Seidenfeld and Schervish (1983). The money pumps are more serious problems for decision-theoretic modelling than the dominance paradox.

3.1. *Indifference between dominating acts*

The basic example is

Example 3.1. *The state space is $S=\mathbb{N}$ and the subjective probability P is the "uniform" distribution λ. The set of consequences is $[-1, +1]$, and $U: [-1, +1]\rightarrow \mathbb{R}$ is a continuous, strictly increasing, expected utility function. If $a_1(n)\downarrow 0$ and $a_1(n)>a_2(n)\geqq 0$ for each $n\in S$, then $U(a_1(n))\downarrow U(0)$ and $U(a_1(n))> U(a_2(n))\geqq U(0)$ for all $n\in S$. But a_1 and a_2 are indifferent because $\int_{\mathbb{N}} U(a_1(n))dP(n)=\int_{\mathbb{N}} U(a_2(n))dP(n)=U(0)$.*[7]

This is pointwise dominance, but at an intuitive level, it is not very much dominance—for any $\varepsilon>0$, a_1 fails to dominate a_2 by even so little as ε on a set having probability 1, $(\forall\varepsilon>0)[P\{n: \varepsilon>a_1(n)>a_2(n)\geqq 0\}=1]$.

An alternative way to understand how little dominance is involved is to note that if a_3 is an act satisfying $a_3(n)\geqq 0$ for all n, then the act a_1-a_3 is indifferent to a_2 if and only if $P\{a_3>\varepsilon\}=0$ for all $\varepsilon>0$. In other words, the amount a_3 that can be taken from an agent with these preferences satisfies both $a_3\geqq 0$ and $\int a_3\,dP=0$. As seen above, with finitely additive probabilities, $a_3(n)>0$ may hold for all n and still $\int a_3\,dP=0$.

To see the role played by the failure of countable additivity, let $E_k=\{|a_1-a_2|<1/k\}$ for $k\in\mathbb{N}$, and $E=\bigcap_k E_k$. If P were countably additive, then $P(E_k)=1$ for all k would imply $P(E)=1$, and E is the event on which a_1 and a_2 are equal. If the state space had some representation for the set E, this paradox would disappear. In any reinterpretation adequate for our purposes, each E_k has a corresponding $\hat{E}_k=\{|\hat{a}_1-\hat{a}_2|\leqq 1/k\}$, and $P(E_k)=\hat{P}(\hat{E}_k)=1$, implying that $\bigcap_k \hat{E}_k=\{|\hat{a}_1-\hat{a}_2|=0\}$ has probability 1, and this is the set on which $\hat{a}_1$ and $\hat{a}_2$ are equal—"not very much dominance" becomes "dominance on a null set".

3.2. *Adams' money pump*

Money pumps are more serious because it seems that a strictly positive amount can be taken from the agents. Both money pumps can be formulated with purely finitely additive probabilities, but this is somewhat less convenient.

Example 3.2 (Adams). *With the state space $S=\mathbb{N}$, let Q be the countably additive probability satisfying $Q\{n\}=2^{-n}$. The subjective probability is $P=(Q+\lambda)/2$ so that $P\{n\}=$*

7. Any non-atomic, finitely additive P would work in this example because it satisfies $P(E)=0$ for any finite E. In describing a variant of his money pump construction, Adams (1962) enumerates the rationals in $[0, 1]$ as $\{q_n: n\in\mathbb{N}\}$, defines a particular non-atomic measure on them, and sets $a_1(q_n)=2^{-n}$, and $a_2(q_n)=0$. In a footnote, Seidenfeld and Schervish (1983) set $a_1(n)=n^{-1}$ and $a_2(n)=0$. Wakker (1993) sets $a_1(n)=n^{-1}$ and $a_2(n)=(n+1)^{-1}$.

$2^{-(n+1)}$ *and* $\sum_{n=1}^{\infty} P\{n\} = \frac{1}{2} < P(\mathbb{N}) = 1$. *The set of consequences is* $[-1, +1]$, *and the expected utility function is* $U(x) = x$. *Fix some* $r \in (\frac{1}{2}, 1)$. *For each* $n \in \mathbb{N}$, *consider the gamble* g_n *that loses* r *if* $B_n = \{n\}$ *occurs, and that pays* $2^{-(n+1)}$ *no matter what occurs,* $g_n(m) = 2^{-(n+1)} - r \cdot 1_{B_n}(m)$ *for* $m \in S$. *This gamble has a positive expected value because* $r < 1$ *and* $P(B_n) = 2^{-(n+1)}$. *By risk neutrality, the agent strictly prefers taking any finite set of these gambles to not taking them. However, for any* $m \in S$, *the payoff to taking all of the gambles simultaneously is*

$$\sum_{n \in \mathbb{N}} g_n(m) = \sum_{n \in \mathbb{N}} (2^{-(n+1)} - r \cdot 1_{B_n}(m)) = \tfrac{1}{2} - r < 0. \tag{4}$$

In other words, the agent will pay a little bit to take each of the gambles *ex ante*, but at the end of this process, will pay to get out of having taken them at each $m \in S$.

After accepting the first N bets, the agent wins $\sum_{i=1}^{N} 2^{-(n+1)}$ in all states of the world, and loses $r < 1$ in the event $E_N = \{1, 2, \ldots, N\}$, a set having probability $\sum_{i=1}^{N} P\{n\} = \sum_{i=1}^{N} 2^{-(n+1)} < \frac{1}{2}$. However, after accepting all of the bets, the agent is in the position of winning $\sum_{n \in \mathbb{N}} 2^{-(n+1)} = \frac{1}{2}$ in all states of the world, and of losing r in the event $E = \bigcup_N E_N = \{1, 2, \ldots\}$, a set having probability 1. Intuitively, $\sum_{n \in \mathbb{N}} P\{n\} = \frac{1}{2}$ means that half of the probability mass is mislaid when enumerating the state space as $\bigcup_N E_N$. The mislaid probability is in the "complement" of E, a set which "should" have probability $\frac{1}{2}$, but which can be empty because P fails countable additivity. If the state space had some representation for the complement of E, this paradox would disappear.

This example is built on the failure of dominated convergence. After accepting the first N bets, the agent's losses are $f_N(m) = -r1_{\{1,\ldots,N\}}(m)$. But

$$\lim_{N \to \infty} \int_{\mathbb{N}} f_N(m)dP(m) = -r/2 > \int_{\mathbb{N}} \lim_{N \to \infty} f_N(m)dP(m) = \int_{\mathbb{N}} -rdP(m) = -r. \tag{5}$$

In other words, the losses increase discontinuously in the limit. By contrast, in any reinterpretation adequate for our purpose, the agent's losses after the first N bets are $\hat{f}_N(m) = -r1_{\{1,\ldots,N\}}(m)$. Here

$$\lim_{N \to \infty} \int_{\hat{\mathbb{N}}} \hat{f}_N(m)d\hat{P}(m) = -r/2 = \int_{\hat{\mathbb{N}}} \lim_{N \to \infty} \hat{f}_N(m)dP(m) = -r \int_{\hat{\mathbb{N}}} 1_{\mathbb{N}}(m)d\hat{P}(m), \tag{6}$$

because $\hat{P}$ is countably additive, and $\hat{P}(\mathbb{N}) = 1/2$. This is possible because $\mathbb{N} \subsetneq \hat{\mathbb{N}}$.

3.3. *Seidenfeld and Schervish's money pump*

This money pump is built on the failure of conglomerability in an example from Dubins (1975).

Example 3.3 (Dubins). *Let* $S = \bigcup\{(i, j) : i \in \mathbb{N}, j = 0, 1\}$, *so that* S *is the union of two copies of the integers, indexed by* $j = 0$ *or* $j = 1$. *Let* $E = \bigcup_i \{(i, 1)\}$ *be the event that* $j = 1$, *and for* $i \in \mathbb{N}$. *Let* $E_i = \{(i, 0), (i, 1)\}$ *so that* $\pi = \{E_1, E_2, \ldots\}$ *is a partition of* S. *Conditional on* E, *suppose that* $P(i, 1) = 1/2(Q + \lambda)(i)$ *where* Q *and* λ *are as in the previous examples. Conditional on* E^c, *suppose that* $P = Q$. *Thus, for any* $i \in \mathbb{N}$, $P(\{(i, 0)\}) = \frac{1}{2} \cdot 2^{-i}$ *and* $P(\{(i, 1)\}) = \frac{1}{4} \cdot 2^{-i}$. *For each* E_i, $P(E|E_i) = \frac{1}{3}$ *even though* $P(E) = \frac{1}{2}$, *so* P *is not conglomerable in* π. *Note that* $\sum_{E_i \in \pi} P(E_i) = \frac{3}{4} < 1$ *even though* π *is a partition.*

Problems arise because a bet on E is a 50:50 affair, while a bet on E conditional on any E_i is a 1:2 affair. Thus, some bets look quite good unconditionally though they look quite bad conditional on each and every event in a countably infinite partition of the state space.

Example 3.4 (Seidenfeld and Schervish). *Suppose that E and π are as in Dubin's example. Let a_1 deliver a consequence worth* 35 *utils in all states while a_2 delivers a consequence worth* 0 *utils if E occurs and* 60 *utils if E does not occur. Because* $35 > \frac{1}{2}(0) + \frac{1}{2}(60) = 30$, $a_2 \prec a_1$. *But $a_1 \prec a_2$ given any E_i because* $40 = \frac{1}{3}(0) + \frac{2}{3}(60) > 35$.

In other words, a person with these preferences would pay to move from a_2 to a_1, and then, conditional on each and every event in a partition of the state space, pay again to move back.

This paradox can also be understood as a failure of dominated convergence. Let D_n be the complement of $\bigcup_{i=1}^{N} E_i$, and let $D = \bigcap_n D_n$. The countable additivity of P would imply that $\lim_n \int 1_{D_n}(m) dP(m) = 1/4 > 0$ would imply that $P(D) = 1/4$. However, with the present state space, the event D is the empty set, giving the appearance of a money pump. If the state space had some representation of the set D, this paradox would also disappear. In any reinterpretation adequate for our purposes, the set $\bigcap_n \hat{D}_n$ has probability 1/4, and conditional on this set, the event $\hat{E}$, on which $\hat{a}_1$ delivers 0 utils, has probability 1.

4. FRAMEWORKS FOR DECISIONS UNDER UNCERTAINTY

4.1. *Bayesian decision frameworks*

One formulation of Savage (1954, 1972) is

Definition 4.1. A Bayesian decision framework is a 4-tuple

$$\mathscr{B} = ((S, \mathscr{S}), (\boldsymbol{C}, \mathscr{C}), \boldsymbol{A}, \boldsymbol{G}).$$

Here (S, $\mathscr{S}$) is a measure space of states, $(\boldsymbol{C}, \mathscr{C})$ is a measure space of consequences, $\boldsymbol{A}$ is a subset (perhaps proper) of the set of $(\mathscr{S}\backslash\mathscr{C})$-measurable functions from S to $\boldsymbol{C}$, and $\boldsymbol{G}$ is a subset (perhaps proper) of the set of $(\mathscr{S}\backslash\mathscr{C})$-measurable functions taking on only finitely many values.[8]

Typically, consequences are denoted by c_1, c_2, or c_3, acts are denoted a_1, a_2, and a_3, and gambles are denoted g_1, g_2, and g_3. Of interest are the complete, reflexive, and transitive preference relations $\preceq$ either on $\boldsymbol{G}$ or on $\boldsymbol{A}$. Of particular interest are those preference relations having SEU representations.

Definition 4.2. A preference relation $\preceq$ on the set of acts $\boldsymbol{A}$ has a subjective expected utility representation for acts (respectively for gambles) if there exists a subjective probability P on $\mathscr{S}$ and a real-valued, $\mathscr{C}$-measurable function U on $\boldsymbol{C}$ such that for all $a \in \boldsymbol{A}$, $\int |U(a)| dP$ *is finite, and for all* $a_1, a_2 \in \boldsymbol{A}$, $a_1 \preceq a_2$ if and only if $\int U(a_1) dP \leqq \int U(a_2) dP$ (respectively, for all $g_1, g_2 \in \boldsymbol{G}$, $g_1 \preceq g_2$ if and only if $\int U(g_1) dP \leqq \int U(g_2) dP$).

8. A measure space is a non-empty set and a σ-field of subsets. To avoid some difficulties, assume that $\boldsymbol{C}$ contains at least two points, and that any σ-field mentioned separates points, if $x_1 \neq x_2$, then there exists a measurable E with $x_1 \in E$ and $x_2 \notin E$. If this assumption is not valid, then simply pass to the set of equivalence classes where, by definition, $x_1 \sim x_2$ if for all measurable E, $1_E(x_1) = 1_E(x_2)$.

4.2. *Savage decision frameworks*

Savage worked with a special class of decision frameworks. He assumed that S is an infinite set, that $\mathcal{S}$ is the set of all subsets of S, and that $\mathcal{C}$ is the set of all subsets of the set of consequences. This implies that the set of Savage acts is the set of *all* functions from S to $\boldsymbol{C}$ and the set of gambles is the set of *all* finite range functions from S to $\boldsymbol{C}$. In a Savage decision framework, if a preference relation $\preceq$ satisfies six postulates, $P1$–6, then it has an SEU representation for gambles. If it also satisfies a seventh postulate, $P7$, then it has an SEU representation for acts. ($P1$–7 are given in Appendix A for ease of reference.)

5. REINTERPRETATIONS OF DECISION FRAMEWORKS

Of primary interest are reinterpretations of Bayesian decision frameworks in which the preference over gambles have SEU representations. However, the reinterpretations provided here are much more broadly applicable.

For these reinterpretations to be adequate for our present purposes, they should provide a copy of the original decision framework and provide a resolution to the paradoxes that arise from failures of countable additivity. This involves three sets of requirements. The first set of requirements concerns the state space, the second set concerns acts (and gambles), and the third set concerns "lining up" the first and the second set of requirements. The "lining up" requirement is that for any bounded continuous v on $\boldsymbol{C}$.

$$\int_S v(a(s))dP(s) = \int_{\hat{S}} v(\hat{a}(\hat{s}))d\hat{P}(\hat{s}). \tag{7}$$

Note that Savage's decision frameworks make no use of continuity assumptions on the expected utility functions U. The last part of this section discusses how the continuity assumptions used here apply, or can be made to apply.

5.1. *A copy of the state space*

For the next four definitions, fix two measure spaces $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$.

Definition 5.1. A mapping Φ from $\mathcal{X}$ into $\mathcal{Y}$ is an isomorphism if it is one-to-one, preserves unions, intersections, and complements, that is, for all $E_1, E_2 \in \mathcal{X}$, $\Phi(E_1 \cup E_2) = \Phi(E_1) \cup \Phi(E_2)$, $\Phi(E_1 \cap E_2) = \Phi(E_1) \cap \Phi(E_2)$, and $\Phi(E_1^c) = (\Phi(E_1))^c$. The class of sets $\Phi(\mathcal{X})$ is a field, and is an isomorphic copy of $\mathcal{X}$ (alternately, Φ is an isomorphism between $\mathcal{X}$ and $\Phi(\mathcal{X})$).

The class $\Phi(\mathcal{X})$ will not be a σ-field in our reinterpretations if it is to resolve the paradoxes. For example, if $\Phi(\bigcap_n D_n) = \bigcap_n \Phi(D_n)$ in the explanation of the Seidenfeld and Schervish money pump, then there would be no non-empty representation of $\bigcap_n \Phi(D_n)$ in the new state space.

As well as providing an isomorphic copy of the collection of subsets on which subjective probabilities are defined, reinterpretations should provide an "isomorphic" copy of the original state space. To this end,

Definition 5.2. Suppose that Φ is an isomorphism from $\mathcal{X}$ into $\mathcal{Y}$. A measurable embedding φ of X in Y is subordinate to Φ if for all $x \in E \in \mathcal{X}$, $\varphi(x) \in \Phi(E)$.

In general, $\{\varphi(s): s \in E\}$ is a strict subset of $\Phi(E)$—the reinterpretations add points to the original space. If P is a finitely additive probability on $\mathscr{X}$ and Φ is an isomorphism from $\mathscr{X}$ to $\mathscr{Y}$, then $\Phi(P)$ can be defined by $\Phi(P)(E) = P(\Phi^{-1}(E))$ for all $E \in \Phi(\mathscr{X})$. Of interest is the existence of countably additive extensions of $\Phi(P)$ to all of $\mathscr{Y}$.

Definition 5.3. An isomorphism Φ from $\mathscr{X}$ to $\mathscr{Y}$ has the unique countably additive extension property if for all finitely additive probabilities P on $\mathscr{X}$, $\Phi(P)$ has a unique countably additive extension from $\Phi(\mathscr{X})$ to $\mathscr{Y}$.

Gathering these properties together,

Definition 5.4. An isomorphism Φ from $\mathscr{X}$ into $\mathscr{Y}$ is adequate for our purposes or simply adequate if it has the unique countably additive extension property and there is an embedding φ subordinate to it.

When Φ is an adequate isomorphism and φ is a given embedding subordinate to Φ, $\Phi(E)$, $\varphi(s)$ and $\Phi(P)$ will be denoted $\hat{E}$, $\hat{s}$, and $\hat{P}$ as convenient.[9]

The new points in Y are $Y \backslash \{\varphi(x): x \in X\}$. In an adequate reinterpretation, these are entities added to X so that the probabilities add up correctly even under countable operations. With P failing countable additivity, $E_n \downarrow E$ in $\mathscr{X}$ does not imply that $P(E_n) \downarrow P(E)$. However, $\bigcap_n \hat{E}_n$ can be a strict superset of $\Phi(\bigcap_n E_n)$ so that $\hat{P}(\hat{E}_n) \downarrow \hat{P}(\bigcap_n \hat{E}_n)$. There is a more sophisticated but parallel interpretation of the extra elements. Think of the events in $\mathscr{X}$ as propositions, so that an event E consists of those states of the world in which a proposition is true. The intersection and union operations can be thought of as the logical operations "&" and "or". The σ-field $\mathscr{Y}$ contains all countable limits of these logical operations. The new points $Y \backslash \{\varphi(x): x \in X\}$ are the states of the world that represent the limits of these operations.

5.2. *A copy of the acts*

A reinterpretation of a Bayesian decision framework should also provide a copy of the acts.

Definition 5.5. Let P be a subjective probability on $(S, \mathscr{S})$ in the Bayesian decision framework $\mathscr{B}$. A reinterpretation of $\mathscr{B}$ for P is a 6-tuple

$$\hat{\mathscr{B}}_P = ((\hat{S}, \hat{\mathscr{S}}), (\boldsymbol{C}, \mathscr{C}), \hat{\boldsymbol{A}}_P, \hat{\boldsymbol{G}}; \Phi, \psi_P).$$

Here $(\hat{S}, \hat{\mathscr{S}})$ is a measure space, Φ is an isomorphism between $\mathscr{S}$ and $\Phi(\mathscr{S}) \subset \hat{\mathscr{S}}$, and ψ_P is a one-to-one, onto mapping from $\boldsymbol{A}$ into the set $\hat{\boldsymbol{A}}_P$ of $\boldsymbol{C}$-valued acts on $(\hat{S}, \hat{\mathscr{S}})$, $a \mapsto \psi_P(a) = \hat{a}$.

By itself, a reinterpretation need not help much, after all, setting $(\hat{S}, \hat{\mathscr{S}}) = (S, \mathscr{S})$ *having* Φ *defined by* $\Phi(E) = \{\varphi(s): s \in E\}$ where φ is a permutation of S, and defining $\psi_P(a)(\varphi(s)) = a(\varphi(s))$ gives a rather useless reinterpretation. What is needed is that Φ have good properties and line up with ψ_P in the correct fashion.

9. Kingman (1967) calls isomorphisms with subordinate embeddings *ramification maps* in his study of finitely additive probabilities in the theory of continuous-time stochastic processes. See Section 8.

5.3. *Lining up the copies*

The next definition provides a broadly useful class of decision frameworks.

Definition 5.6. A Bayesian decision framework $\mathscr{B}$ is csm (complete, separable, and metric) (respectively compact) if the set of consequences, $\boldsymbol{C}$, is a complete, separable, metric (respectively a compact metric) space,[10] and $\mathscr{C}$ is the Borel σ-field on $\boldsymbol{C}$. An act a is nearly compactly supported for the subjective probability P if for all $\varepsilon>0$, there exists a compact set $K_\varepsilon \subset \boldsymbol{C}$ such that $P\{a\in K_\varepsilon\}>1-\varepsilon$. A csm $\mathscr{B}$ is nearly compactly supported for the subjective probability P if all acts are nearly compactly supported for P.

If $\mathscr{B}$ is compact, then it is nearly compactly supported for all P, though in general the set of nearly compactly supported acts depends on P. When consequences are monetary, the near compactness assumption is that for every act a, and every $\varepsilon>0$, there exists an $N_\varepsilon \in \mathbb{R}$ such that $P\{a\in[-N_\varepsilon, +N_\varepsilon]\}>1-\varepsilon$. If P were countably additive, this would follow either from the observation that $\mathbb{R}$ is a complete, separable metric space (so all countably additive probabilities are nearly compactly supported), or from the observation that $\mathbb{R}=\bigcup_{N\in\mathbb{N}}[-N, +N]$. Thus, acts being nearly compactly supported is necessary for the existence of reinterpretations in many interesting cases.

The copies "line up" if the integrals of bounded continuous functions are preserved.

Definition 5.7. For any subjective probability P on $(S, \mathscr{S})$ in a csm $\mathscr{B}$, the reinterpretation $\hat{\mathscr{B}}_P$ is adequate for P if Φ is an adequate isomorphism from $\mathscr{S}$ into $\hat{\mathscr{S}}$, and the mapping $a \mapsto \psi_P(a)=\hat{a}$ has the property that for $a\in \boldsymbol{A}$, and for all bounded continuous functions $v\colon \boldsymbol{C}\to\mathbb{R}$,

$$\int_S v(a)dP=\int_{\hat{S}} v(\hat{a})d\hat{P}. \tag{8}$$

Equation (8) guarantees that any SEU preferences $\precsim$ over a set of nearly compactly supported acts having a representation with a continuous expected utility function U is duplicated by defining $\hat{a}_1 \hat{\precsim} \hat{a}_2$ if and only if $a_1\precsim a_2$. The class of preferences that is duplicated in an adequate reinterpretation is much broader than the set of SEU preferences. It contains all state independent preferences that do not distinguish between weak*-equivalent distributions on the set of consequences.

The finitely additive probabilities μ and μ' on $(\boldsymbol{C}, \mathscr{C})$ are *weak*-equivalent* if for any bounded continuous v, $\int_C v d\mu=\int_C v d\mu'$. By the Reisz representation theorem (for finite, countably additive measures on complete, separable metric spaces), every finitely additive, nearly compactly supported probability μ is weak*-equivalent to a unique countably additive probability ca(μ). Equation (8) can be restated as saying that ca$(P_a)=\hat{P}_{\hat{a}}$ where P_a is the distribution on $\boldsymbol{C}$ induced by the nearly compactly supported a and $\hat{P}_{\hat{a}}$ is the corresponding distribution after an adequate reinterpretation. Suppose that $\precsim$ is a preference ordering over acts that is *state independent*, that is, a preference ordering that depends *only* on the distributions induced by the acts. Suppose further that it does not distinguish between weak*-equivalent distributions. The implication of (8) is that $\precsim$ is duplicated in any adequate reinterpretation simply by defining $\hat{a}_1 \hat{\precsim} \hat{a}_2$ if and only if $a_1\precsim a_2$.

10. The real line with the usual topology is a complete, separable metric (csm) space. Intervals $[a, b]$ have all of these properties and are also compact. Complete, separable metric spaces are sometimes called Polish spaces. Many of the subsequent results are true in greater generality than the restriction to csm Bayesian decision frameworks implies, and this is noted.

The class of preference orderings that depend only on distributions over consequences and do not distinguish between weak*-equivalent distributions includes Machina's (1982) generalized expected utility preferences. It also includes the "non-additive" preferences represented by $a_1 \precsim a_2$ if and only if $\int u dg(F_{U(a_1)}(u)) \leqq \int u dg(F_{U(a_2)}(u))$ where U is a continuous, real-valued function on $\boldsymbol{C}$, g is a continuous, increasing function from [0, 1] onto [0, 1], and $F_{U(a)}$ is the c.d.f. of the distribution $U(P_a)$.[11] Because these classes of preferences contain SEU preferences as a special case, they trivially contain money pumps and dominance puzzles if the subjective probabilities fail countable additivity. Adequate reinterpretations resolve those parts of the money pumps due to this failure, and not those parts due to a failure of the "linearity" of SEU preferences.

5.4. *Properties of adequate reinterpretations*

A crucial property of adequate reinterpretations is that they exist.

Theorem 5.1. *For any* csm $\mathscr{B}$ *which is nearly compactly supported for the subjective probability P, a reinterpretation adequate for P exists.*

A simplification of the proof delivers the following.

Corollary 5.2. *For any compact* $\mathscr{B}$, *there is an adequate reinterpretation that is independent of P.*[12]

Gambles (simple acts) have special properties in reinterpretations—note the lack of restrictions on $\mathscr{B}$ and v in the following.

Corollary 5.3. *For any* $\mathscr{B}$, *there is a reinterpretation* $\hat{\mathscr{B}}$ *that is independent of P such that for all finitely additive probabilities P on* $\mathscr{S}$, *all gambles* $g \in \boldsymbol{G}$, *and for all functions* $v: \boldsymbol{C} \to \mathbb{R}$,

$$\int v(g)dP = \int v(\hat{g})d\hat{P}. \tag{9}$$

Equation (9) guarantees the *equality*, not merely the weak*-equivalence, of P_g and $\hat{P}_{\hat{g}}$ for gamble g. Thus, any state independent preference ordering over simple acts is duplicated by defining $\hat{g}_1 \hat{\precsim} \hat{g}_2$ if and only if $g_1 \precsim g_2$.

A potentially inconvenient aspect of using adequate reinterpretations is that the preferences $\hat{\precsim}$ are defined only on $\hat{\boldsymbol{A}}_P$, generally a strict subset of $M((\hat{S}, \hat{\mathscr{S}}); \boldsymbol{C})$ (the measurable functions from $\hat{S}$ to $\boldsymbol{C}$). There may be measurable functions on $\hat{S}$ that induce distributions

11. The additive case is $g(r) \equiv r$. Quiggin (1993) contains an extended treatment of this class of generalized expected utility preferences. Allais (1953, p. 510, 512) suggests using systematic changes in the distribution over consequences, verbally identifying some currently popular properties the function g. He also (p. 513) suggests studying preferences that can be represented by non-linear functions of the entire distribution of $U(P_a)$. These weak*-continuous preferences are either Machina (1982) preferences or are arbitrarily close to this class of preferences. It was Machina's work that first showed that such an approach can deliver a workable theory of choice under uncertainty. See also Allais and Hagen (1979).

12. All that is needed for Theorem 5.1 and Corollary 5.2 is that the space $\boldsymbol{C}$ be a Hausdorff topological space. One proof of these results uses the Loeb (1975) spaces $({}^*S, \sigma({}^*\mathscr{S}))$, where $\hat{P}$ is the Loeb measure derived from ${}^*P$. When $\mathscr{B}$ is not compact, $\hat{a} := {}^{\circ}({}^*a)$ will have to be modified on a set having $\hat{P}$-measure 0. The near compactness assumption is necessary for this step to work.

over consequences very different than the set of distributions induced in $\mathscr{B}$. The next result says that this cannot happen for non-atomic P, and can barely happen for general P.

For any class of (measurable) functions, H from X to Y, and any probability Q on X, $H(Q)$ denotes the set of probabilities Q' on Y of the form $Q'(E) = Q\{h^{-1}(E)\}$ for some $h \in H$.

Theorem 5.4. *Suppose that $\mathscr{B}$ is* csm, *and that A is the set of all nearly compactly supported acts for a subjective probability P. If P is non-atomic, then $\hat{A}_P(\hat{P}) = M(\hat{P})$ in any reinterpretation adequate for P where $M = M((\hat{S}, \hat{\mathscr{S}}); C)$. For general P, $M(\hat{P})$ is the variation norm closure of $A_P(\hat{P})$.*[13]

5.5. *Savage's lack of assumed structure*

One of the beauties of Savage decision frameworks is the lack of assumed structure. Savage makes no assumptions on the space of consequences. By contrast, the very definition of adequate reinterpretations requires at least a topological structure on the space of consequences and a corresponding measure theoretic structure. This contrast is not as large at it may appear. The following example demonstrates the major difficulty that must be overcome in generalizing adequate reinterpretations to arbitrary Savage decision frameworks.

Example 5.1. *Suppose that $S = \mathbb{N}$, that $C = \mathbb{R}$, and that the subjective probability is (say) the "uniform" distribution $P = \lambda$. Let $a_1(s) \equiv s$. If preferences are over all acts and are state independent, they must cover the distribution $P_{a_1} = \lambda$ on C. However, Section* 2.6 *shows that "draws" made according to a distribution that fails countable additivity are carried on points not contained in the original space.*

It is possible to systematically embed S in $\hat{S}$ so as to provide points that carry the subjective probability as a countably additive probability. In a similar fashion, it is also possible to systematically embed C in a space $\hat{C}$ so as to provide points that carry all possible P_a as countably additive probabilities. One such procedure, and there are many, compactifies C in such a fashion that *all* bounded functions, v, on C correspond uniquely to a continuous $\hat{v}$ on the compactified space, $\hat{C}$. This can be interpreted as the observation that Savage *implicitly* worked with continuous expected utility functions just as he implicitly worked with a measurable structure. The topology in which his expected utility functions are continuous is the finest possible, namely the set of all subsets of C. Generalized adequate reinterpretations require adding points to the topological space C so as to provide carriers for finitely additive probabilities. With the appropriate generalization of weak*-equivalence, the results above carry through.[14]

13. The variation norm distance between two probabilities μ and ν is given by $\sup_E |\mu(E) - \nu(E)|$. Appendix C contains an example of an atomic P in a compact $\mathscr{B}$ for which $M(\hat{P})$ strictly contains $\hat{A}(\hat{P})$.

14. On the (strong) advice of the editor and judging from the (strong) tendency of my friends to change the topic when faced with the details of generalized adequate reinterpretations, no more will be said here!

6. ADEQUATELY REINTERPRETING THE PARADOXES

6.1. *The dominance puzzle*

The starting point is

Lemma 6.1. *Suppose that* $\mathscr{B}$ *is* csm,[15] P *is a subjective probability, and that* A *is the set of all nearly compactly supported, measurable functions. If* $P\{a \in F\} = 1$ *for some (measurable) set* $F \subset C$, *then* $\hat{P}$*-almost everywhere,* $\hat{a}$ *takes its value in the closure of* F *in any reinterpretation adequate for* P.

Recall that Example 3.1 concerned two acts, a_1 and a_2, defined on the space of integers and satisfying $\lim_{n \to \infty} a_1(n) = 0$ and $a_1(n) > a_2(n) \geqq 0$ for all $n \in \mathbb{N}$. Because $P(A) = 0$ for all finite A, for all $k \in \mathbb{N}$, $P(E_k) = 1$ where $E_k = \{n \in \mathbb{N}: |a_1(n) - a_2(n)| < 1/k\}$. Because the set of consequences is compact and U is continuous, Theorem 5.1 delivers an adequate reinterpretation in which U is the expected utility function representing $\hat{\precsim}$ on $\hat{A}$. By definition then, $\hat{P}(\hat{E}_k) = 1$. Lemma 6.1 implies that up to at most a set of $\hat{P}$-measure 0, $\hat{E}_k = \{n \in \hat{\mathbb{N}}: |\hat{a}_1(n) - \hat{a}_2(n)| \leqq 1/k\}$ (provided that $|a_1(n) - a_2(n)|$ is itself an act in the original $\mathscr{B}$). The countable additivity of $\hat{P}$ then implies that $\hat{P}(E') = 1$ where $E' = \bigcap_k \hat{E}_k$. But $\hat{a}_1$ and $\hat{a}_2$ are equal on E'. In other words, in any adequate reinterpretation, the dominance puzzle reduces to dominance on a null set.

6.2. *Adams' money pump*

The gambles in Example 3.2 are of the form $g_n(m) = 2^{-(n+1)} - r1_{B_n}(m)$ for $m \in \mathbb{N}$ and $n \in \mathbb{N}$. They are mapped to $\hat{g}_n(m) = 2^{-(n+1)} - r1_{\hat{B}_n}(m)$ where $B_n = \{n\}$ for $m \in \hat{\mathbb{N}}$ and $n \in \mathbb{N}$. By Theorem 5.1, the agent is risk neutral in any adequate reinterpretation. After taking the first N gambles in any adequate reinterpretation, the agent's position is $\hat{p}_N = \sum_{n=1}^{N} \hat{g}_n$, and $\hat{p}_N \uparrow \frac{1}{2}1_{\hat{\mathbb{N}}} - r1_{\mathbb{N}}$. (In this last expression, it is important to note the difference between $\hat{\mathbb{N}}$ and $\mathbb{N}$.) Because $P\{n\} = \hat{P}\{n\} = 2^{-(n+1)}$ and $\hat{P}$ is countably additive, $\hat{P}(\mathbb{N}) = \frac{1}{2} < \hat{P}(\hat{\mathbb{N}})$. Thus, $\int \hat{p}_N d\hat{P} = \frac{1}{2} - r\hat{P}(\mathbb{N}) = \frac{1}{2} - \frac{1}{2}r$, exactly the limit of the expected positions from taking the gambles sequentially. The mislaid probability in the original, inadequate analysis is carried by the set $\hat{\mathbb{N}} \backslash \mathbb{N}$, and there is no paradox.

6.3. *Seidenfeld and Schervish's money pump*

The starting point in Example 3.4 is the observation that the partition $\pi = \{E_1, E_2, \ldots\}$ satisfies $\sum_i P(E_i) < 1$, so that $\sum_i \hat{P}(\hat{E}_i) < 1$. Because $\hat{P}$ in any adequate reinterpretation is countably additive, this means that $\bigcup_i \hat{E}_i$ is a strict subset of $\hat{S}$. In particular, $\hat{\pi} = \{\hat{E}_1, \hat{E}_2, \ldots\}$ is *not* a partition of $\hat{S}$—evaluating a gamble conditional on each and every $\hat{E}_i$ is not a complete analysis. In particular, conditional on the complement of $\bigcup_i \hat{E}_i$, the event E has probability 1.

Suppose that the three consequences in Example 3.4 are monetary, and satisfy $U(c_1) = 0$, $U(c_2) = 60$, and $U(c_3) = 35$, and that U is continuous, bounded, and strictly increasing. Theorem 5.1 applies, and the continuity and monotonicity of U imply that there exist

15. This lemma requires only that the space of consequences be a normal topological space.

quantities of money, $r_1, r_2 > 0$ such that

$$U(c_3 - r_1) > \tfrac{1}{2}U(c_1) + \tfrac{1}{2}U(c_2), \tag{10}$$

and

$$\tfrac{1}{3}U(c_1 - (r_1 + r_2)) + \tfrac{2}{3}U(c_2 - (r_1 + r_2)) > U(c_3 - r_1). \tag{11}$$

If the initial situation is the gamble g described by c_1 if E and c_2 if E^c, then its reinterpretation is the gamble $\hat{g}$ given by c_1 if $\hat{E}$ and c_2 if $\hat{E}^c$. Inequality (10) implies that this agent is strictly willing to buy insurance, that is, to pay r_1 in each state of the world in return for being guaranteed c_3. Now, conditional on each and every E_i (respectively each and every $\hat{E}_i$), inequality (11) implies that this agent is strictly willing to pay r_2 in each state $s \in E_i$ (respectively in each state $\hat{s} \in \hat{E}_i$) in return for being guaranteed $c_1 - r_1$ if E and $c_2 - r_1$ if E^c. (This follows from Bayes' Law and the definition of an isomorphism.) In the original money pump, one concludes that, because π is a partition, the agent was always happy to make the trades that reduced the initial situation of c_1 if E and c_2 if E^c to the terminal situation of $c_1 - (r_1 + r_2)$ if E and $c_2 - (r_1 + r_2)$ if E^c. By contrast, in any adequate reinterpretation of the money pump, the agent is changed from the initial situation of c_1 if $\hat{E}$ and c_2 if $\hat{E}^c$ to the terminal situation only partially described by $c_1 - (r_1 + r_2)$ if $\hat{E} \cap (\bigcup_i \hat{E}_i)$ and $c_2 - (r_1 + r_2)$ if $\hat{E}^c \cap (\bigcup_i \hat{E}_i)$. The description of the terminal situation is not complete until it specifies what happens on the non-empty set $D' = (\bigcup_i \hat{E}_i)^c$.

6.4. *Summary*

The point of view taken here is that the original decision frameworks for the money pumps and the dominance paradoxes are simply not adequate. Countably infinite constructions require countably additive probabilities. The paradoxes use acts whose salient properties are described by countable limit operations, but do not use a countably additive probability. When this lack is corrected by reinterpretations that provide mirror copies of the original decision frameworks, the money pumps and the dominance paradoxes disappear because it is no longer possible to ignore sets of states having positive probability.

7. MINIMAL REINTERPRETATIONS

The previous two sections have shown that it is possible to reinterpret finitely additive Bayesian decision frameworks in an adequate fashion, and so resolve the paradoxes. There is, however, a loose end. This is due to the new state spaces $\hat{S}$ being so much larger than the original state space S. The expansion of a state space S in a decision framework $\mathscr{B}$ to larger state space $\hat{S}$ gives rise to a decision framework having many new subjective probabilities. In principle, the new decision framework, $\hat{\mathscr{B}}$, may give rise to distributions unlike any on the original state space.

Example 7.1. *An adequate reinterpretation of a Bayesian decision framework based on the two-point state space* $(S, \mathscr{S}) = (\{0, 1\}, 2^{\{0,1\}})$ *can be based on the (rather larger) space* $(\hat{S}, \hat{\mathscr{S}}) = ([0, 1], \mathscr{A})$ *where* $\mathscr{A}$ *is the usual Borel* σ*-field, with the mapping* $\varphi(s) = \hat{s} = s$, *and* $\Phi(E) = \{\varphi(s) : s \in E\}$. *Applied to acts, the embedding requires* $\hat{a}(\hat{s}) = a(s)$ *for* $\hat{s} = 0, 1$, *and makes no requirements for the other* $\hat{s} \in (0, 1)$.

The new decision framework in this Example is clearly adequate, but it is so much larger than it need be that it would never be used. One possible countably additive subjective probability of an agent with the state space $\hat{S}$ is the uniform distribution. With the right acts, this gives rise to all of the possible distributions over any Polish (csm) space of consequences. By contrast, distributions induced by subjective probabilities on the original space must have two point supports. An intuitive criterion is that the enlarged state space allow only countably additive probabilities that arise as the image of finitely additive probabilities on the original space.

Definition 7.1. *An adequate isomorphism Φ from $\mathcal{S}$ to $\hat{\mathcal{S}}$ is minimal if every countably additive probability on $\hat{\mathcal{S}}$ is of the form $\Phi(Q)$ for some finitely additive probability Q on $\mathcal{S}$. An adequate reinterpretation $\hat{\mathcal{B}}$ is minimal if Φ is minimal.*

Requiring minimality implies that the new state space in a reinterpretation is a Stone space.[16] A measurable isomorphism of two measure spaces is a one-to-one, onto measurable mapping with measurable inverse. From the measure theoretic point of view, two measurably isomorphic spaces are indistinguishable. Kingman (1967, Theorem 5) proves that if Φ is a minimal isomorphism from $\mathcal{S}$ to $\hat{\mathcal{S}}$, then $(\hat{S}, \hat{\mathcal{S}})$ is measurably isomorphic to the Stone space for the Boolean algebra of bounded, measurable, real-valued functions on $(S, \mathcal{S})$. This explains

Theorem 7.1. *For any* csm *$\mathcal{B}$ which is nearly compactly supported for the subjective probability P, a minimal adequate reinterpretation for P exists. If $\mathcal{B}$ is compact and the set of acts is the set of all measurable functions, then there is a unique minimal reinterpretation which is adequate for all P.*

The previous section showed that adequate reinterpretations resolve the paradoxes that arise from the failure of countable additivity, all of this in an isomorphic copy of the original framework. This section has shown that there are adequate reinterpretations that add nothing to the set of phenomena being modelled. The cost of this minimality is that the Stone space is rather peculiar. If the original state space was (say) $S=[0, 1]$ with $\mathcal{S}$ the set of Borel subsets of S, then the space $\hat{S}$ is known as the Stone space for $L_\infty[0, 1]$. This is a very large, compact space in which every open set is also closed. Many of the convenient and comfortable features of [0, 1] are lost. Beyond the observation that reinterpretations resolve the paradoxes and minimal reinterpretations are still reinterpretations, there are two responses to this cost:

1. it isn't that big—any analysis of choice under uncertainty that depends in a crucial fashion on special properties of the state space seems a bit misguided. If a result in choice theory is *only* true with a state space equivalent to [0, 1] with the Borels, then there may be problems with the result.
2. it can be avoided—the space $\hat{S}$ can be avoided by using a *non*-minimal reinterpretation. For example, the Loeb (1975) spaces used in the proof of Theorem 5.1 have the same first order logic properties as the space S.

16. The original references are Stone (1937, 1947–48), and these spaces are covered in many texts, e.g. Dunford and Schwartz (1957), Semadeni (1971), or Sikorski (1969). For completeness, a simple construction of Stone spaces is given in Appendix B.

8. METHODOLOGICAL REFLECTIONS

The authors of the paradoxes presented above wrote to point out that there are conflicts between finitely additive subjective probabilities and desirable properties for a theory of decisions under uncertainty. One summary of their work is that countably infinite constructions require countably additive probabilities. This paper proposes to resolve the conflict by reinterpreting decision frameworks, by identifying the finitely additive probabilities with countably additive probabilities on larger state spaces. Addition of new states in this fashion rules out ignoring positive probability sets of states by clever countable operations. The argument here is that the reinterpreted decision framework is so close to the original one that, at the very least, if countable constructions are to be allowed, the new state space should be regarded as a "truer" version of the state space. This is a claim with a contentious history.

Skyrms (1995) provides a wonderfully clear discussion of the links between the different kinds of additivity and metaphysical arguments about models of quantities. Of particular importance in this work are the results relating the additivity of subjective probabilities and whether or not money pumps exist. Many of the tensions that arise between the different philosophical considerations can be solved by use of a class of nonstandard probability spaces that are called star-finite (e.g. Anderson (1982)). These spaces can also provide adequate reinterpretations.

Villegas' (1964) work also uses the minimal reinterpretations provided by a Stone space. Villegas provides an axiom, monotone continuity, which, if satisfied, guarantees that the subjective probability is countably additive. This is clearly not compatible with Savage's modelling choices. However, Villegas also argued that, because any finitely additive probability "can be extended to a monotonely continuous qualitative probability" (in the terminology of this paper, can be identified with its countably additive extension to the appropriate Stone space), "there is no loss in generality if we consider only qualitative probabilities which are monotonely continuous". This paper has argued that the reinterpretations are different than the original frameworks, although they are very close. In particular, (1) the state spaces are more complicated, (2) measurability requirements are necessary, it is no longer that case that the agents can assign a probability to each and every set of states,[17] and (3) the decision-theoretic paradoxes that result from the failure of countable additivity are resolved.

Minimal reinterpretations have appeared in statistical decision theory. Le Cam (1986, Ch. 1.6, pp 11–15) provides several characterizations of the appropriate Stone space for statistical decision theory, but seems to regard the whole question of state spaces as more of a nuisance than anything else. For Le Cam's decision theory, probabilities are continuous linear functions on lattices of real-valued functions, and it is these structures that are in the foreground.

The new points in the minimal reinterpretations have played a parallel role in Arrow Impossibility Theorems for models with infinite sets of agents. Kirman and Sondermann (1972) and Hansson (1976) independently built on Fishburn's (1970) Arrow *possibility* example when the set of agents is infinite. Suppose that $\mathscr{F}$ is a free ultrafilter in an infinite set of agents.[18] Define a social order by $x \succ y$ if and only if $x \succ_i y$ for every i in some set $I \in \mathscr{F}$. This gives a social choice rule that satisfies all of Arrow's assumptions, and the

17. Skyrms (1993) and (1995) contains discussions of the issues of interpreting state spaces and non-measurable "events".

18. A class $\mathscr{F} \subset 2^X$ of subsets of an infinite set X is called a free ultrafilter if $A \in \mathscr{F}$ implies that A is not finite, $A, B \in \mathscr{F}$ implies that $A \cap B \in \mathscr{F}$, $A \in \mathscr{F}$ and $B \supseteq A$ implies that $B \in \mathscr{F}$, and for all $A \in 2^X$, either $A \in \mathscr{F}$ or $A^c \in \mathscr{F}$.

authors show that this is the only way to satisfy Arrow's assumptions. As Kirman and Sondermann noted, the free ultrafilters are equivalent to the new points in a Stone space of agents, and they identify these agents as "invisible dictators". These new points are individuals that represent the intersection of increasingly small (generalized) sequences of sets that decrease to the empty set in the original space.[19]

The point of view advanced here, that the original frameworks are not adequate for the problems being considered, most closely resembles Kingman's (1967) analysis of finitely additive probabilities in the study of continuous-time stochastic processes and Harris, Stinchcombe, and Zame's (1995) analysis of finitely-additive mixed strategies in games. As in the Arrow possibility theorems, it is the emptiness of intersections that shouldn't be empty that drives both sets of arguments.

Harris, Stinchcombe, and Zame (1995) argue that the failure of equilibrium arguments due to a lack of ideal elements is not generally interesting. Starting from the minimal structure of games, they examine equilibrium arguments with ideal elements, essentially replacing inadequate strategy spaces with versions of the $\hat{S}$ of this paper. What makes their analysis complicated is that the products of finitely additive probabilities fail the conclusions of Fubini's theorem. They show that adding ideal elements to strategy sets directly parallels the present addition of ideal elements to state space only for those games in which Fubini's theorem holds.

Kingman (1967) begins with the observation that countable additivity is violated if and only if there is a sequence of sets $E_n \downarrow \emptyset$ with $\lim P(E_n) = \delta > 0$. Kingman argues that $\delta > 0$ means that it is the space that "is defective and should be regarded as a subset of some larger space in which the required (points) exists". For example, he proves that any process, be it a Brownian motion, a pure jump process, or one with everywhere non-measurable sample paths, has a representation as a non-countably additive probability on the set of polynomial paths. Because the polynomials are inadequate to represent these phenomena, the probabilities must fail countable additivity. He then proceeds to minimal isomorphisms as a way of adding ideal elements to the state space.[20]

This addition of states in order to guarantee countable additivity runs entirely counter to de Finetti's (1972, 1974, 1975) point of view. Regarding Kingman's work, he writes (1975, p. 353),

> The basic idea is the possibility of stretching the interpretation in such a way as to be able to attribute the "missing" probability in the partition to new fictitious entities in order that everything adds up properly. In some cases, in order to salvage countable additivity, it is even claimed that the new entities are not fictitious, but real.

As I read it, de Finetti's argument begins with the observation that it is only possible to sample a continuous-time stochastic process at a finite collection of times. Because non-measurable paths or jumps can only be observed with an infinite set of samples, one can argue that these "continuous" properties are the "fictitious" ones.

de Finetti also notes (esp. 1974, pp 229–231) that finitely additive probabilities can be extended, perhaps in many ways, to any class of sets, while countably additive probabilities, when they can be extended at all, extend uniquely. He argues for the desirability of

19. This whole issue is more thoroughly covered in Armstrong (1980, 1985). Hansson objects to this interpretation of ultrafilters for two reasons. First, there may not be a pre-defined, meaningful topology on the space of agents so that compactification may not mean much. (Kirman and Sondermann are quite careful with topological issues in their interpretation.) Second, if there is a topology, it may not be Hausdorff, and this leads to equivalence classes of "dictators".

20. The looser notion of an adequate reinterpretation would have also served Kingman's purposes.

many extensions, asking for a case-by-case choice between finitely and countably additive probabilities, and disdaining "a preconceived preference for that which yields a unique and elegant answer *even when the exact answer should instead be 'any value lying between these limits'*" (italics in the original). There are serious philosophical difficulties with countable additivity is the interpretation of the non-measurable sets—these are "events" or propositions to which an agent cannot assign a degree of likelihood. de Finetti's argument is that this difficulty implies that finite additivity is the correct choice.

What the money pumps show is that beyond being "unique and elegant" (and this is no small virtue), the countably additive version of the theory is the right one for the case of decisions under uncertainty if our agents should not be exploitable because of some " 'missing' probability". Of course, this argument is susceptible to the counter that the money pumps and dominance examples require infinite constructions, and are therefore just as "fictitious" as (say) jump processes. That being said, the closeness of the reinterpretations convinces this author that the "fictitious" states are better thought of as being initially unobserved,[21] though no less "real" for that.

As economists, we do not want to build models of markets with agents that can be money pumped. With subjective probabilities that fail countable additivity, there are money pumps for SEU preferences or any of their generalizations. For all of these preferences, the "fix" proposed here also works. The fix involves adding extra states, in this case and in many related ones, the states are easy to interpret.

APPENDIX A. SAVAGE'S POSTULATES

The definitions used in stating the Postulates are included.

*P*1. $\preceq$ is complete and transitive.

*D*1. $a_1 \preceq a_2$ given $B \subset S$, if and only if $a_1' \preceq a_2'$ for every a_1' and a_2' that agree with a_1 and a_2, respectively, on B and with each other on B^c and $a_2' \preceq a_1'$ either for all such pairs or for none.

*P*2. For every a_1, a_2, and $B \subset S$, either $a_1 \preceq a_2$ given B or $a_2 \preceq a_1$ given B.

*D*2. $c_1 \preceq c_2$ if and only if $a_1 \preceq a_2$ when $a_1(s) = c_1$ and $a_2(s) = c_2$ for every $s \in S$.

*D*3. $B \subset S$ is null if and only if for all a_1, a_2, $a_1 \preceq a_2$ given B.

*P*3. For every non-null $B \subset S$, if $a_1(s) = c_1$ and $a_2(s) = c_2$ for every $s \in B$, then $a_1 \preceq a_2$ given B if and only if $c_1 \preceq c_2$.

*D*4. For A, $B \subset S$, $A \leqq B$ if and only if $a_A \preceq a_B$ or $c_1 \preceq c_2$ for every a_A, a_B, c_1, c_2 such that $a_A(s) = c_1$ for $s \in A$, $a_A(s) = c_2$ for $s \notin A$, $a_B(s) = c_1$ for $s \in B$, $a_B(s) = c_2$ for $s \notin B$.

*P*4. For every A, $B \subset S$, $A \leqq B$ or $B \leqq A$.

*P*5. It is false that for every a_1 and a_2, $a_1 \preceq a_2$.

*P*6. Suppose that $a_2 \succ a_3$. Then, for every $c_1 \in C$ there is a finite partition of S such that, if a_2' agrees with a_2 and a_3' agrees with a_3 except on an arbitrary element of the partition, a_2' and a_3' being equal to c_1 there, then $a_2' \succ a_3$ and $a_2 \succ a_3'$.

If the preference relation $\preceq$ also satisfies the following Postulate, then it has a SEU representation for acts.

*D*5. $a_1 \preceq c_1$ given $B \subset S$ (respectively $c_1 \preceq a_1$ given B) if and only if $a_1 \preceq a_2$ given B (respectively $a_2 \preceq a_1$ given B) when $a_2(s) = c_1$ for every $s \in S$.

*P*7. If $a_1 \preceq a_2(s)$ (respectively $a_2(s) \preceq a_1$) given $B \subset S$ for every $s \in B$, then $a_1 \preceq a_2$ given B (respectively $a_2 \preceq a_1$ given B).

21. Thanks to an anonymous referee for this felicitous phrase.

APPENDIX B. A CONSTRUCTION OF THE STONE SPACE

Let $(X, \mathscr{X})$ be a measure space, and let M_b denote the set of bounded, measurable, real-valued functions on X. Each $x \in X$ can be uniquely identified with the infinite vector $v(x) := (f(x))_{f \in M_b}$ in the space $X' := \times_{f \in M_b}[\inf_{x \in X} f(x), \sup_{x \in X} f(x)]$. With the product topology, X' is compact (by Tychonoff's theorem). The Stone space for L_∞ is the compact closure, $\bar{V}$, in X' of the set $V := \{v(x): x \in X\}$. In this space, each $f \in M_b$ has a unique continuous extension from V to $\bar{V}$, and all continuous functions on $\bar{V}$ are extensions of $f \in M_b$.

APPENDIX C. PROOFS

All non-standard constructions are assumed to be in a polysaturated extension of a superstructure containing S as a bounded set (see e.g. Hurd and Loeb (1985) or Lindstrøm (1988) for accessible introductions to non-standard analysis).

Proof of Theorem 5.1. Set $\hat{S} = {}^*S$, set $\hat{\mathscr{S}} = \sigma({}^*\mathscr{S})$, and let $\hat{P}$ be the Loeb (1975) measure derived by extending the finitely additive ${}^{\circ *}P$ from ${}^*\mathscr{S}$ to $\sigma({}^*\mathscr{S})$. The isomorphism Φ is defined by $\Phi(E) = {}^*E$.

We first check that Φ is adequate. Define φ by $\varphi(s) = {}^*s$. By transfer, Φ is an isomorphism and φ is subordinate to Φ. The basic result of Loeb (1975) is that Φ has the unique countably additive extension property, so that Φ is adequate as needed.

We must now define ψ_P. To this end, let c_0 be an arbitrary point in C. For any act a, let $E_n = a^{-1}(K_n)$ where $K_n \subset C$ is compact and satisfies $P\{a \in K_n\} > 1 - 1/n$. By the assumption that acts are nearly compactly supported, such a K_n exists. For $s \in \bigcup_n {}^*E_n$, define $\psi_P(a)(s) = \hat{a}(s) = {}^{\circ *}a(s)$. For $s \notin \bigcup_n {}^*E_n$, set $\hat{a}(s) = c_0$. The function $\hat{a}$ is measurable, and if v is bounded and continuous, then

$$\int v(a)dP = {}^*\!\int {}^*v({}^*a)d{}^*P \simeq \int {}^{\circ *}v({}^*a)d\hat{P} = \int v({}^{\circ *}a)d\hat{P} = \int v(\hat{a})d\hat{P}.$$

The first equality is definitional, the second follows because $v(a)$ is bounded and measurable, the third because v is continuous and a is nearly compactly supported, and the fourth because ψ_P modifies ${}^*a$ on the complement of $\bigcup_n {}^*E_n$, a set of $\hat{P}$-measure 0. ||

Proof of Corollary 5.2. If C is compact, then set $K_n \equiv C$ in the construction used in the proof of Theorem 5.1, so that $\hat{a}$ is independent of P. ||

Proof of Corollary 5.3. Endow the space C with the discrete topology. Any gamble is automatically nearly compactly supported as finite sets are compact, and any functions v is continuous. Bounding v outside the range of a gamble if necessary, the above construction of Φ and ψ_P works for gambles with no need for modification on sets of measure 0. ||

Proof of Lemma 6.1. Suppose, for the purposes of contradiction, that $\hat{a}$ does not take its values $\hat{P}$-almost everywhere in $\bar{F}$, the closure of F. Let $\hat{P}_{\hat{a}}$ denote the distribution on C induced by $\hat{a}$ acting on $\hat{P}$. Because $\hat{P}$ is countably additive and C is normal (and every metric space is normal), if $\hat{P}(\bar{F}) < 1$, then there exists a compact, hence closed, F_2, disjoint from $\bar{F}$ and such that $\hat{P}(F) > 0$. By the compactness of F_2 and Urysohn's characterization of normal spaces, there exists a continuous function v satisfying $0 \leq v \leq 1$, $v(F_2) = 1$, and $v(\bar{F}) = 0$. But the definition of an adequate reinterpretation implies that

$$0 = \int v(a)dP = \int v(\hat{a})d\hat{P} > 0.$$

a contradiction that completes the proof. ||

Proof of Theorem 5.4. Suppose that P is non-atomic, i.e. for all $E \in \mathscr{S}$ and $0 < r < 1$, there exists $E' \subset E$, $E' \in \mathscr{S}$ such that $P(E') = rP(E)$ (by Armstrong and Prikry (1981) this is equivalent to non-atomicity). Pick an arbitrary $m \in M$. We must show that for some nearly compactly supported act $a \in A$, $\hat{P}_{\hat{a}} = \hat{P}_m$.

Because all countably additive probabilities on a Polish set of consequences are nearly compactly supported, there exists a sequence of compact $K_n \subset C$ such that $\hat{P}(m^{-1}(K_n)) > 1 - 1/n$ and $K_n \subset K_{n+1}$. Because each K_n is compact, there exists a sequence of measurable, finite partitions of C, $(D_{n,k})_{k=1}^{K(n)+1}$ such that (i) $\bigcup_{k=1}^{K(n)} D_{n,k} = K_n$ (so that $D_{n,K(n)+1} = C \backslash K_n$), (ii) $\max\{\text{diameter } D_{n,k}: k \leq K(n)\} < 1/n$, and, (iii) for all n, $(D_{n+1,k})_{k=1}^{K(n+1)}$ restricted to K_n refines $(D_{n,k})_{k=1}^{K(n)}$. The next step is to produce an increasingly fine class of partitions of S, $(E_{n,k})_{k=1}^{K(n)+1}$, to match up with $m^{-1}(D_{n,k})$.

Because P is non-atomic, there exists a parallel sequence of finite, measurable partitions of S, $(E_{n,k})_{k=1}^{K(n)+1}$ such that (i′) $P(E_{n,k}) = \hat{P}(m^{-1}(D_{n,k}))$, (iii′), for all n, restricted to $E_n := \bigcup_{k=1}^{K(n)} E_{n,k}$, the partition $(E_{n+1,k})_{k=1}^{K(n+1)}$ refines the partition $(E_{n,k})_{k=1}^{K(n)}$. For each (n, k) pair with $k \leqq K(n)$, pick $c_{n,k} \in D_{n,k}$ and fix and arbitrary $c_0 \in C$. For $n \in \mathbb{N}$, define $a_n(s) = c_{n,k}$ if $s \in E_{n,k}$ and $a_n(s) = c_0$ otherwise. For all $s \in \bigcup_n E_n$, $a_n(s)$ is a Cauchy sequence and so converges because C is complete. For $s \notin \bigcup_n E_n$, $a_n(s) \equiv c_0$. Hence $a_n \to a$ for all $s \in S$ and some measurable $a \in A$. For any $\varepsilon > 0$ and for all n sufficiently large, the construction of a_n guarantees that the Prohorov distance between $\hat{P}_{\hat{a}_n}$ and $\hat{P}_m$ is less than ε. Thus, for $\hat{a}$, the Prohorov distance between $\hat{P}_{\hat{a}}$ and $\hat{P}_m$ is 0.

For compact metric sets of consequences, the same construction works with increasingly fine partitions of all of C of the form $(D_{n,k})_{k=1}^{K(n)}$ with maximum diameter going to 0, and the construction is therefore independent of P.

For the general case, the equality of variation norm closures follows from dividing P into its atomic and its non-atomic parts. The first half of this proof covered the non-atomic part, the atomic part is immediate. ‖

The following is an example of an Atomic P in a compact $\mathscr{B}$ for which $M(\hat{P})$ strictly contains $\hat{A}(\hat{P})$.

Example C.1. *Define P on $\mathbb{N}$ by $P = \frac{1}{3}Q_1 + \frac{2}{3}Q_2$ where $Q_1(\{n\}) = 2^{-n}$, and Q_2 is a $\{0, 1\}$-valued, purely finitely additive measure on $\mathbb{N}$ such that $Q_2(A) = 0$ for any finite set A. Enumerate the rationals in $C = [0, 1]$ as $\{q_n : n \in \mathbb{N}\}$. Let $a(n) = q_n$. Note that the act a has a unique continuous extension, $\hat{a}$, to $\hat{\mathbb{N}}$, the Stone space for the Boolean algebra of bounded, real-valued functions on $\mathbb{N}$. The measure $\hat{Q}_2$ is point mass on some point $n' \in \hat{\mathbb{N}}$. This means that $P_r := \hat{a}(\hat{P})$ puts mass $\frac{1}{3}2^{-n}$ on the n-th rational q_n, and puts mass $\frac{2}{3}$ on some $r \in [0, 1]$. This is the r uniquely determined by $r = \hat{a}(n')$. Any continuous function f on $\hat{\mathbb{N}}$ with the property that $f(\hat{P})(q_n) = \frac{1}{3}2^{-n}$ must satisfy $f(n') = r$ because $\mathbb{N}$ is dense in $\hat{\mathbb{N}}$, and continuous functions are determined by their values on dense sets. Thus, no act b on $\mathbb{N}$ can have $\hat{b}(\hat{P}) = P_s$, $s \neq r$, the measure which mass $\frac{1}{3}2^{-n}$ on the n-th rational q_n, and puts mass $\frac{2}{3}$ on s.*

Proof of Theorem 7.1. Define the equivalence relation $\sim$ on $^*S$ by $s \sim t$ if for all measurable, bounded, real-valued f on S, $^*f(s) \simeq {}^*f(t)$. By Anderson (1982), $^*S/\sim$ with the weakest topology making each of the $^{\circ *}f$ continuous is the Stone space for the Boolean algebra of bounded measurable functions on $(S, \mathscr{S})$. The construction of ψ_P in the proof of Theorem 5.1 then gives a measurable $\hat{a}$ because the indicator function, 1_{E_n}, of each $E_n = a^{-1}(K_n)$ is bounded and measurable. Further, the interpretation is adequate for P because taking equivalence classes does not affect the integrals because $v(a)$ is bounded and measurable. The independence of P for compact $\mathscr{B}$ follows from Corollary 5.2. ‖

Acknowledgements. Bob Anderson, Mark Machina, Joel Sobel, Hal White, the editor, and three anonymous referees helped immensely with perspective on, questions about, references for, and organization of this paper. They are not to blame for the remaining deficiencies, they prevented worse.

REFERENCES

ADAMS, E. (1962), "On Rational Betting Systems", *Archiv für Mathematische Logik und Grundlagenforschung*, **6**, 7–18 and 112–128.

ALLAIS, M. (1953), "Le Comportment de l'Homme Rationnel Devant le Risque: Critique des Postulates et Axiomes de l'Ecole Americaine", *Econometrica*, **21**, 503–546 (English translation in Allais and Hagen (1979)).

ALLAIS, M. and HAGEN, O. (1979) *Expected Utility Hypotheses and the Allais Paradox: Contemporary Discussions of Decisions Under Uncertainty with Allais' Rejoinder* (Dordrecht: D. Reidel Publishing Company).

ANDERSON, R. (1982), "Star-finite Representations of Measure Spaces", *Transactions of the American Mathematical Society*, **271**, 667–687.

ARMSTRONG, T. E. (1980), "Arrow's Theorem with Restricted Coalition Algebras", *Journal of Mathematical Economics*, **7**, 55–75.

ARMSTRONG, T. E. (1985), "Precisely Dictatorial Social Welfare Functions", *Journal of Mathematical Economics*, **14**, 57–59.

ARMSTRONG, T. E. (1990), "Conglomerability of Probability Measures on Boolean Algebras", *Journal of Mathematical Analysis and Applications*, **150**, 335–358.

ARMSTRONG, T. and PRIKRY, K. (1981), "Liapounoff's Theorem for Non-Atomic, Bounded, Finitely Additive, Finite Dimensional Vector Valued Measures", *Transactions of the American Mathematical Society*, **266**, 499–514.

AUMANN, R. and SHAPLEY, L. (1974) *Values of Non-Atomic Games* (Princeton: Princeton University Press).

BILLINGSLEY, P. (1979) *Probability and Measure* (New York: John Wiley & Sons).

de FINETTI, B. (1972) *Probability, Induction and Statistics* (New York: J. Wiley & Sons).
de FINETTI, B. (1974) *Theory of Probability, Vol. 1* (translated by A. Machi and A. Smith) (New York: Wiley).
de FINETTI, B. (1975) *Theory of Probability, Vol. 2* (translated by A. Machi and A. Smith) (New York: Wiley).
DUBINS, L. (1975), "Finitely Additive Conditional Probabilities, Conglomerability and Disintegrations", *Annals of Probability*, **3**, 89–99.
DUNFORD, N. and SCHWARTZ, J. (1957) *Linear Operators, Part I: General Theory* (New York: J. Wiley & Sons).
FISHBURN, P. (1970), "Arrow's Impossibility Theorem: Concise Proof and Infinite Voters", *Journal of Economic Theory* **2**, 103–106.
HANSSON, B. (1976), "The Existence of Group Preference Functions", *Public Choice*, **28**, 75–79.
HARRIS, C., STINCHCOMBE, M. and ZAME, W. (1995), "Equilibrium Existence for Infinite Games: The Nearly Compact and Continuous Case" (Working Paper, Department of Economics, University of Texas at Austin).
HEATH, D. and SUDDERTH, W. (1989), "Coherent Inference from Improper Priors and from Finitely Additive Priors", *Annals of Statistics*, **17**, 907–919.
HURD, A. E. and LOEB, P. A. (1985) *An Introduction to Nonstandard Real Analysis* (New York: Academic Press).
KARNI, E. (1993), "Subjective Expected Utility Theory with State-dependent Preferences", *Journal of Economic Theory*, **60**, 428–439.
KINGMAN, J. F. C. (1967), "Additive Set Functions and the Theory of Probability", *Proceedings of the Cambridge Philosophical Society*, **63**, 767–775.
KREPS, D. (1988) *Notes on the Theory of Choice* (Boulder: Westview Press).
KIRMAN, A. P. and SONDERMANN, D. (1972), "Arrow's Theorem, Many Agents, and Invisible Dictators", *Journal of Economic Theory*, **5**, 267–277.
LE CAM, L. (1986) *Asymptotic Methods in Statistical Decision Theory* (New York: Springer-Verlag).
LINDSTRØM, T. (1988), "An Invitation to Nonstandard Analysis", in N. Cutland (ed.), *Nonstandard Analysis and its Applications* (New York: Cambridge University Press), 1–105.
LOEB, P. (1975), "Conversion from Nonstandard to Standard Measure Spaces and Applications in Probability Theory", *Transactions of the American Mathematical Society*, **211**, 113–122.
MACHINA, M. (1982), "'Expected Utility' Analysis without the Independence Axiom", *Econometrica*, **50**, 277–323.
MACHINA, M. and SCHMEIDLER, D. (1992), "A More Robust Definition of Subjective Probability", *Econometrica*, **60**, 745–780.
QUIGGIN, J. (1993) *Generalized Expected Utility Theory: The Rank-dependent Model* (Boston: Kluwer Academic Publishers).
ROYDEN (1968) *Real Analysis* (New York: The Macmillan Company).
SAVAGE, L. J. (1954, 1972) *The Foundations of Statistics* (New York: Wiley (Second edition 1972. New York: Dover)).
SCHMEIDLER, D. (1989), "Subjective Probability and Expected Utility without Additivity", *Econometrica*, **57**, 571–587.
SEIDENFELD, T. and SCHERVISH, M. (1983), "A Conflict Between Finite Additivity and Avoiding Dutch Book", *Philosophy of Science*, **50**, 398–412.
SEIDENFELD, T., SCHERVISH, M. and KADANE, J. B. (1984), "The Extent of Non-conglomerability of Finitely Additive Probabilities", *Wahrscheinlichkeits Theorie Verwandte Gebiete*, **66**, 205–226.
SEMADENI, Z. (1971) *Banach Spaces of Continuous Functions* (Warsaw: Polish Scientific Publishers).
SIERPINSKI, W. (1956) *Hypothése du Continu* (New York: Chelsea, 2nd ed. (first ed. 1934)).
SIKORSKI, R. (1969) *Boolean Algebras* (Berlin: Springer-Verlag).
SKYRMS, B. (1993), "Logical Atoms and Combinatorial Possibility", *The Journal of Philosophy*, **XC**, 219–232.
SKYRMS, B. (1995), "Strict Coherence, Sigma Coherence, and the Metaphysics of Quantity", *Philosophical Studies*, **77**, 39–55.
STONE, M. H. (1937), "Applications of the Theory of Boolean Rings to General Topology", *Transactions of the American Mathematical Society*, **41**, 375–481.
STONE, M. H. (1947-48), "The Generalized Weierstrass Approximation Theorem", *Mathematics Magazine*, **21**, 167–184, 237–254.
ULAM, S. (1930), "Sur Masstheorie in der allgemeinenen Mengenlehre", *Fundamenta Mathematicae*, **16**, 140–150.
VILLEGAS, C. (1964), "On Quantitative Probability σ-Algebras", *Annals of Mathematical Statistics*, **35**, 1787–1796.
WAKKER, P. (1993), "Savage's Axioms Usually Imply Violation of Strict Stochastic Dominance", *Review of Economic Studies*, **60**, 487–493.