DOMINATING COUNTABLY MANY FORECASTS

By M. J. Schervish, Teddy Seidenfeld and J. B. Kadane

Carnegie Mellon University

We investigate differences between a simple Dominance Principle applied to sums of *fair prices* for variables and dominance applied to sums of *forecasts* for variables scored by proper scoring rules. In particular, we consider differences when fair prices and forecasts correspond to finitely additive expectations and dominance is applied with infinitely many prices and/or forecasts.

1. Introduction. The requirement that preferences are *coherent* aims to make rigorous the idea that elementary restrictions on rational preferences entail that personal probabilities satisfy the axioms of mathematical probability. This use of coherence as a justification of personal probability is very well illustrated by de Finetti's (1974) approach to the foundations of probability. De Finetti distinguished two senses of coherence: coherence$_1$ and coherence$_2$. Coherence$_1$ requires that probabilistic forecasts for random variables (he calls them previsions) do not lead to a finite set of fair contracts that, together, are uniformly dominated by abstaining. Coherence$_2$ requires that a finite set of probabilistic forecasts cannot be uniformly dominated under Brier (squared error) score by a rival set of forecasts. He showed that these two senses of coherence are equivalent in the following sense. Each version of coherence results in using the expectation of a random variable as its forecast. Moreover, these expectations are based on a finitely additive probability without requiring that personal probability is countably additive. [In Appendix A, we explain what we mean by expectations with respect to finitely additive probabilities. These are similar in many ways, but not identical to integrals in the sense of Dunford and Schwartz (1958), Chapter III.] Schervish, Seidenfeld and Kadane (2009) extended this equivalence to include a large class of strictly proper scoring rules (not just Brier score) but for events only. The corresponding extension to general random variables is included in the supplemental article [Schervish, Seidenfeld and Kadane (2014)]. Here, we refer to the extended sense of coherence$_2$ as coherence$_3$.

We investigate asymmetries between coherence$_1$ and coherence$_3$ reflecting differences between cases where personal probabilities are countably additive and where personal probabilities are finitely (but not countably) additive. We give conditions where coherence$_3$ may be applied to assessing countably many forecasts

at once, but where coherence$_1$ cannot be applied to combining infinitely many fair contracts. Also, we study conditional forecasts given elements of a partition π, where the conditional forecasts are based on the conditional probabilities given elements of π. Each coherence criterion is violated by combining infinitely many conditional forecasts when those conditional forecasts are not conglomerable (see Definition 7) in the partition π. Neither criterion is violated by combining infinitely many conditional forecasts when conditional expectations satisfy the law of total previsions (see Definition 8) in π.

2. Results of de Finetti.   Coherence of preference, as de Finetti [(1974), Chapter 3] formulates it, is the criterion that a rational decision maker respects *uniform (strict) dominance*. In Section 2.1, we explain the version of the Dominance Principle that de Finetti uses. In Section 2.2, we review de Finetti's two versions of coherence, with a focus on how preferences based on a finitely additive probability are coherent.

2.1. *Dominance.*   Let Ω be a set. The elements of Ω will be called *states* and denoted ω. Random variables are real-valued functions with domain Ω, which we denote with capital letters. Let I index a set of options. Consider a hypothetical decision problem $\mathcal{O}$ specified by a set of exclusive options $\mathcal{O} = \{O_i : i \in I\}$. Each option O_i is a random variable with the following interpretation: If ω is the state which occurs, then $O_i(\omega)$ denotes the decision maker's loss (negative of cardinal utility) for choosing option O_i. The values of O_i (for all $i \in I$) are defined up to a common positive affine transformation.

DEFINITION 1.   Let O_i and O_j be two options from $\mathcal{O}$. If there exists an $\varepsilon > 0$ such that for each $\omega \in \Omega$, $O_j(\omega) > O_i(\omega) + \varepsilon$, then option O_i *uniformly strictly dominates* O_j. If, for each ω, $O_j(\omega) > O_i(\omega)$, we say that O_i *simply dominates* O_j.

Uniform strict dominance is clearly stricter than simple dominance. As we explain, next, in order to permit preferences based on maximizing finitely (and not necessarily countably) additive expectations, de Finetti used the following Dominance Principle, rather than some other more familiar concepts of admissibility, for example, simple dominance. There are additional ways to define dominance, which we discuss further in Section 6.

DOMINANCE PRINCIPLE: Let O_i and O_j be options in $\mathcal{O}$. If O_i uniformly (strictly) dominates O_j, then O_j is an *inadmissible* choice from $\mathcal{O}$.

2.2. *Coherence$_1$ and coherence$_2$.*   De Finetti [(1974), Chapter 3] formulated two criteria of *coherence* that are based on the Dominance Principle. Throughout this paper, we follow the convention of identifying events with their indicator functions.

DEFINITION 2.   A *conditional prevision* (or *conditional forecast*) $P(X|H)$ for a random variable X given a nonempty event H is a fair price for buying and selling X in the sense that, for all real α, the option that costs the agent $\alpha H[X - P(X|H)]$ is considered fair. [We call $P(X|\Omega)$ an *unconditional prevision* and denote it $P(X)$.] A collection $\{P(X_i|H_i) : i \in I\}$ of such conditional forecasts is *coherent*$_1$ if, for every finite subset $\{i_1, \ldots, i_n\} \subseteq I$ and all real $\alpha_1, \ldots, \alpha_n$, there exists no $\varepsilon > 0$ such that

$$\sum_{j=1}^{n} \alpha_j H_{i_j}(\omega)\big[X_{i_j}(\omega) - P(X_{i_j}|H_{i_j})\big] \geq \varepsilon$$

for all $\omega \in \Omega$.

A collection of conditional forecasts is *coherent*$_2$ if no sum of finitely many (Brier score) penalties can be uniformly strictly dominated in the partition of states by the sum of penalties from a rival set of forecasts for the same random variables. That is, for every finite subset $\{i_1, \ldots, i_n\} \subseteq I$, all alternative forecasts $q_{i_1}, \ldots, q_{i_n}$, and all positive $\alpha_1, \ldots, \alpha_n$, there is no $\varepsilon > 0$ such that

$$\sum_{j=1}^{n} \alpha_j H_{i_j}(\omega)\big[X_{i_j}(\omega) - P(X_{i_j}|H_{i_j})\big]^2 \geq \sum_{j=1}^{n} \alpha_j H_{i_j}(\omega)\big[X_{i_j}(\omega) - q_{i_j}\big]^2 + \varepsilon$$

for all ω.

De Finetti [(1974), pages 88–89] proved that a decision maker who wishes to be both coherent$_1$ and coherent$_2$ must choose the same forecasts for both purposes. He also proved that the decision maker's coherent$_1$ forecasts are represented by a finitely additive personal probability, $P(\cdot)$, in the sense of Definition 3 below.

If $P(H) = 0$, then coherence$_1$ and coherence$_2$ place no restrictions on $P(X|H)$ for bounded X. Nevertheless, it is possible and useful to make certain intuitive assumptions about conditional forecasts given events with 0 probability. In particular, Theorems 3 and 4 of this paper assume that $P(\cdot|H)$ is a finitely additive expectation (in the sense of Definition 10 in Appendix A) satisfying $P(X|H) = P(HX|H)$ for all H and X. This assumption holds whenever $P(H) > 0$, and it captures the idea that $P(\cdot|H)$ is concentrated on H. De Finetti [(1975), Appendix 16] introduces an axiom that places a similar requirement on conditional previsions. See Levi (1980), Section 5.6, and Regazzini (1987) for other ways to augment the coherence criteria of Definition 2 in order to satisfy these added requirements on conditional previsions given a null event. Rather than adding such requirements to the definition of coherence, we prefer that individual agents who wish to adopt them do so as explicit additional assumptions. Example 2 in the supplemental article [Schervish, Seidenfeld and Kadane (2014)] illustrates our reason for such a preference. In this way, our definition of coherence is slightly weaker than that of de Finetti.

As an aside, the meaning of conditional expected value in the finitely-additive theory differs from its meaning in the countably-additive theory in this one major regard: In the finitely-additive theory a conditional expectation can be specified given an arbitrary nonempty event, regardless of whether that event has positive probability. A conditional expectation of a bounded random variable given an event with zero probability is not defined uniquely in terms of unconditional expectations, but Dubins (1975) shows that, in the finitely additive theory, conditional expectations can be defined on the set of bounded random variables so that they are finitely additive expectations. In the countably-additive theory, conditional expectation is defined twice: given events with positive probability and given σ-fields. The two definitions match in a well-defined way, and both provide uniquely defined conditional expectations in terms of unconditional expectations.

DEFINITION 3. A probability $P(\cdot)$ is *finitely additive* provided that, when events F and G are disjoint, that is, when $F \cap G = \varnothing$, then $P(F \cup G) = P(F) + P(G)$. A probability is *countably additive* provided that when F_i ($i = 1, \ldots$) is a denumerable sequence of pairwise disjoint events, that is, when $F_i \cap F_j = \varnothing$ if $i \neq j$, then $P(\bigcup_{i=1}^{\infty} F_i) = \sum_{i=1}^{\infty} P(F_i)$. We call a probability P *merely finitely additive* when P is finitely but not countably additive. Likewise, then its P-expectations are merely finitely additive.

For each pair X and Y of random variables with finite previsions (expectations), $P(X + Y) = P(X) + P(Y)$. For countably additive expectations and countably many random variables $\{X_i\}_{i=1}^{\infty}$, conditions under which $P(\sum_{i=1}^{\infty} X_i) = \sum_{i=1}^{\infty} P(X_i)$ can be derived from various theorems such as the monotone convergence theorem, the dominated convergence theorem, Fubini's theorem and Tonelli's theorem.

De Finetti (1981) recognized that coherence$_2$ (but not coherence$_1$) provided an incentive compatible solution to the problem of mechanism design for eliciting a coherent set of personal probabilities. Specifically, Brier score is a *strictly proper scoring rule*, as defined here.

DEFINITION 4. A *scoring rule* for coherent forecasts of a random variable X is a real-valued loss function g with two real arguments: a value of the random variable and a forecast q. Let $\mathcal{P}_g$ be the collection of probability distributions such that $P(X)$ is finite and $P[g(X, q)]$ is finite for at least one q. We say that g is *proper* if, for every probability $P \in \mathcal{P}_g$, $P[g(X, q)]$ is minimized (as a function of q) by $q = P(X)$. If, in addition, only the quantity $q = P(X)$ minimizes expected score, then the scoring rule is *strictly proper*.

The following trivial result connects proper scoring rules with conditional distributions.

PROPOSITION 1. *If H is a nonempty event and $P(\cdot|H)$ is a probability distribution then $P[g(X, q)|H]$ is (uniquely) minimized by $q = P(X|H)$ if g is (strictly) proper.*

Some authors reserve the qualification strictly proper for scoring rules that are designed to elicit an entire distribution, rather than just the mean of a distribution. [See Gneiting (2011a), who calls the latter kind *strictly consistent*.] For the remainder of this paper, we follow the language of Definition 4, which matches the usage in Gneiting (2011b).

We present some background on strictly proper scoring rules in Section 3. Section 4 gives our main results. We discuss propriety of scoring rules for infinitely many forecasts in Section 5.

3. Background on strictly proper scoring rules. In this section, we introduce a large class of strictly proper scoring rules that we use as generalizations of Brier score. Associated with this class, we introduce a third coherence concept that generalizes coherence$_2$.

DEFINITION 5. Let $\mathcal{C}$ be a class of strictly proper scoring rules. Let $\{(X_i, H_i) : i \in I\}$ be a collection of random variable/nonempty event pairs with corresponding conditional forecasts $\{p_i : i \in I\}$. The forecasts are *coherent$_3$ relative to $\mathcal{C}$* if, for every finite subset $\{i_j : j = 1, \ldots, n\} \subseteq I$, every set of scoring rules $\{g_j\}_{j=1}^n \subseteq \mathcal{C}$, and every set $\{q_j\}_{j=1}^n$ of alternative forecasts, there is no $\varepsilon > 0$ such that

$$\sum_{j=1}^n H_{i_j}(\omega) g_j(X_{i_j}(\omega), p_{i_j}) \geq \sum_{j=1}^n H_{i_j}(\omega) g_j(X_{i_j}(\omega), q_j) + \varepsilon$$

for all ω. That is, no sum of finitely many scores can be uniformly strictly dominated by the sum of scores from rival forecasts.

Coherence$_2$ is the special case of coherence$_3$ in which $\mathcal{C}$ consists solely of Brier score. The supplemental article [Schervish, Seidenfeld and Kadane (2014)] includes a proof that, if $\mathcal{C}$ consists of strictly proper scoring rules of the form (1) below, then coherence$_3$ relative to $\mathcal{C}$ is equivalent to coherence$_1$.

The general form of scoring rule that we will consider is

$$(1) \qquad g(x, q) = \begin{cases} \displaystyle\int_x^q (v - x) \, d\lambda(v), & \text{if } x \leq q, \\ \displaystyle\int_q^x (x - v) \, d\lambda(v), & \text{if } x > q, \end{cases}$$

where λ is a measure that is mutually absolutely continuous with Lebesgue measure and is finite on every bounded interval. It is helpful to rewrite (1) as

$$(2) \qquad g(x, q) = \int_q^x (x - v) \, d\lambda(v),$$

using the convention that an integral whose limits are in the wrong order equals the negative of the integral with the limits in the correct order. Another interesting way to rewrite (1), using the same convention, is

$$g(x, q) = \lambda((q, x))[x - r(x, q, \lambda)], \tag{3}$$

where, for all a and b,

$$r(a, b, \lambda) = \frac{\int_a^b v \, d\lambda(v)}{\lambda((a, b))}. \tag{4}$$

An immediate consequence of (3) is that, if p and q are real numbers, then

$$g(x, q) - g(x, p) = \lambda((q, p))[x - r(q, p, \lambda)]. \tag{5}$$

The form (1) is suggested by equation (4.3) of Savage (1971). Each such scoring rule is finite, nonnegative and continuous as a function of (x, q). If we wanted to consider only countably additive distributions, we could use a larger class of scoring rules by allowing λ to be an infinite measure supported on a bounded interval (c_1, c_2). But this relaxation would allow functions g that are not strictly proper for natural classes of finitely additive distributions. Example 1 below illustrates this point. Lemma 1 justifies the use of (1) as the form of our scoring rules. The proofs of all results in the body of the paper are given in Appendix B.

LEMMA 1. *Let g be a scoring rule of the form* (1). *Then g is strictly proper.*

It follows from (3) that, if λ is a probability measure with finite mean, then $\mathcal{P}_g$ from Definition 4 is the class of all finitely additive distributions with finite mean because $\lambda((q, x))$ and $\lambda((q, x))r(q, x, \lambda)$ are both bounded functions of q and x. Even if λ is not a finite measure, (5) implies that, if $P[g(X, q_0)]$ is finite, then $h(x, p) = g(x, p) - g(x, q_0)$ is linear in x so that $h(x, p)$ is also strictly proper with $\mathcal{P}_h$ equal to the class of all probabilities with finite mean. For example, if $g(x, p) = (x - p)^2$, namely Brier score, then $\mathcal{P}_g$ is the set of distributions with finite second moment. However, $h(x, p) = (x - p)^2 - x^2$ has $\mathcal{P}_h$ equal to the class of all probabilities with finite mean.

Let $f(\cdot)$ denote the Radon–Nikodym derivative of λ with respect to Lebesgue measure. Some familiar examples of strictly proper scoring rules are recovered by setting f equal to specific functions. Brier score corresponds to $f(v) \equiv 2$. Logarithmic score on the interval (c_1, c_2) corresponds to $f(v) = (c_2 - c_1)/[(c_2 - v)(v - c_1)]$, but the corresponding measure is infinite on (c_1, c_2). Hence, logarithmic score is not of the form (1). In addition, if g is this logarithmic score, then $\mathcal{P}_g$ does not include all finitely additive distributions that take values in the bounded interval (c_1, c_2), as the following example illustrates.

EXAMPLE 1. Let X be a random variable whose entire distribution is agglutinated at c_1 from above. That is, let $P(X > c_1) = 1$ and $P(X < c_1 + \varepsilon) = 1$ for all $\varepsilon > 0$. Let g be the logarithmic scoring rule that uses $f(v)$ from above. Then $P(X) = c_1$, but $g(X(\omega), c_1) = \infty$ for all ω, which could not have finite mean even if we tried to extend the definition of random variables to allow them to assume infinite values. On the other hand, for $c_1 < q < c_2$, the mean of $g(X, q)$ is $\log[(c_2 - c_1)/(c_2 - q)] > 0$, which decreases to 0 as q decreases to c_1, and is always finite. So, $\mathcal{P}_g$ is nonempty but does not contain P.

Some of our results rely on one or another condition that prevents the λ measures that determine the scoring rules from either being too heavily concentrated on small sets or from being too different from each other.

DEFINITION 6. Let $\mathcal{C} = \{g_i : i \in I\}$ be a collection of strictly proper scoring rules of the form (1) with corresponding measures $\{\lambda_i : i \in I\}$.

(i) Suppose that, for every $\varepsilon \geq 0$, there exists $\delta_\varepsilon > 0$ such that for all $i \in I$ and all real $a < b$, $\lambda_i((a, b)) > \varepsilon$ implies $a + \delta_\varepsilon \leq r(a, b, \lambda_i) \leq b - \delta_\varepsilon$. Then we say that the collection $\mathcal{C}$ satisfies the *uniform spread condition*.

(ii) Suppose that, for every $\varepsilon > 0$ and every $i \in I$, there exists $\gamma_{i,\varepsilon} > 0$ such that for all $j \in I$ and all real $a < b$, $\lambda_i((a, b)) \geq \varepsilon$ implies $\lambda_j((a, b)) \geq \gamma_{i,\varepsilon}$. Then we say that the collection $\mathcal{C}$ satisfies the *uniform similarity condition*.

The $r(a, b, \lambda)$ in (4) can be thought of as the mean of the probability measure on the interval (a, b) obtained by normalizing λ on the interval. The uniform spread condition insures that, the λ measures are spread out enough to keep the means of the normalized measures on intervals far enough away from both endpoints.

The next result gives sufficient conditions for both the uniform similarity and uniform spread conditions. It is easy to see that the conditions are logically independent of each other.

LEMMA 2. *Let* $\mathcal{C} = \{g_i : i \in I\}$ *be a collection of strictly proper scoring rules of the form* (1) *with corresponding measures* $\{\lambda_i : i \in I\}$ *and corresponding Radon–Nikodym derivatives* $\{f_i : i \in i\}$ *with respect to Lebesgue measure.*

(i) *Assume that there exists* $U < \infty$ *such that* $f_i(v) \leq U$, *for all* v *and all* $i \in I$. *Then* $\mathcal{C}$ *satisfies the uniform spread condition.*

(ii) *Assume that for every* $i \in I$, *there exists* $L_i > 0$ *such that* $f_j(v)/f_i(v) \geq L_i$, *for all* v *and all* $j \in I$. *Then* $\mathcal{C}$ *satisfies the uniform similarity condition.*

As an example, suppose that each λ_i is $\alpha_i > 0$ times Lebesgue measure. If the α_i are bounded above, then $\mathcal{C}$ satisfies the uniform spread condition. If the α_i are bounded away from 0, then $\mathcal{C}$ satisfies the uniform similarity condition. These sets of measures correspond to multiples of Brier score. There are collections that

satisfy the uniform spread condition without satisfying the conditions of part (i) of Lemma 2. For example, let $f(v) = |v|^{-1/2}/2$ which is not bounded above. For this f, we have $\lambda((a, b)) = |\sqrt{|b|} - \sqrt{|a|}|$ if $0 \notin (a, b)$, and $\lambda((a, b)) = \sqrt{|a|} + \sqrt{|b|}$ if $0 \in (a, b)$. So $\lambda((a, b))^2$ is no larger than two times the distance between a and b. Also, $r(a, b, \lambda)$ is always at least $1/3$ of the way from both a and b. We can add the corresponding scoring rule to any class that already satisfies the uniform spread condition by (if necessary) lowering δ_ε to $\varepsilon^2/6$.

4. Extensions to countably many options. In Section 4.1, we investigate when each sense of coherence can be extended to allow combining countably many forecasts into a single act by summing together their individual outcomes. In Section 4.2, we introduce the concept of *conditional* forecasts and present results about the combination of countably many coherent conditional forecasts.

4.1. *Dominance for countably many forecasts.* Let $\{X_i\}_{i=1}^\infty$ be a countable set of random variables with corresponding coherent$_1$ unconditional previsions $\{p_i\}_{i=1}^\infty$. Let $\{\alpha_i\}_{i=1}^\infty$ be a sequence of real numbers. The decision maker's net loss in state ω, from adding the individual losses from the fair options $\alpha_i[X_i(\omega) - p_i]$ is

$$(6) \qquad \sum_{i=1}^\infty \alpha_i[X_i(\omega) - p_i].$$

Similarly, if the agent's prevision p_i for X_i is scored by the strictly proper scoring rule g_i for each i, the total score in each state ω equals

$$\sum_{i=1}^\infty g_i(X_i(\omega), p_i).$$

We assume that each of the two series above are convergent for all $\omega \in \Omega$.

EXAMPLE 2 (Combining countably many forecasts). De Finetti [(1972), page 91] noted that when the decision maker's personal probability is merely finitely additive, she/he cannot always accept as fair the countable sum (6) determined by coherent$_1$ forecasts. That sum may be uniformly dominated by abstaining. Let $\Omega = \{\omega_i\}_{i=1}^\infty$ be a countable state space. Let W_i be the indicator function for state $\omega_i : W_i(\omega) = 1$ if $\omega = \omega_i$ and $W_i(\omega) = 0$ if $\omega \neq \omega_i$. Consider a collection of merely finitely additive coherent$_1$ forecasts $P(W_i) = p_i \geq 0$ where $\sum_{i=1}^\infty p_i = c < 1$. So $P(\cdot)$ is not countably additive. With $\alpha_i = 1$, for all i, the loss from combining these infinitely many forecasts into a single option is uniformly positive,

$$\sum_{i=1}^\infty \alpha_i[W_i(\omega) - p_i] = (1 - c) > 0.$$

Hence, the decision maker's alternative to abstain, with constant loss 0, uniformly strictly dominates this infinite combination of fair options.

If, on the other hand, the decision maker's personal probability P is countably additive, then $c = 1$. For arbitrary $\{\alpha_i\}_{i=1}^{\infty}$ such that $d = \sum_{i=1}^{\infty} \alpha_i p_i$ is defined and finite, the sum of losses is

$$(7) \qquad \sum_{i=1}^{\infty} \alpha_i \big[W_i(\omega) - p_i \big] = \alpha_{i(\omega)} - d,$$

where $i(\omega)$ is the unique i such that $W_i(\omega) = 1$. Because $c = 1$, there is at least one $\alpha_i \leq d$ and at least one $\alpha_i \geq d$, hence (7) must be nonpositive for at least one i, and abstaining does not uniformly strictly dominate.

Next, we focus on the parallel question whether a coherent$_3$ set of forecasts remains undominated when strictly proper scores for countably many forecasts are summed together. Some conditions will be needed in order to avoid $\infty - \infty$ arising in the calculations, and these are stated precisely in the theorems. The principal difference between dominance for infinite sums of forecasts and dominance for infinite sums of strictly proper scores is expressed by the following result.

THEOREM 1. *Let $\mathcal{C}$ be a collection of strictly proper scoring rules of the form* (1) *that satisfies the uniform spread condition. Let P be a coherent$_3$ prevision defined over a collection $\mathcal{D}$ of random variables that contains all of the random variables mentioned in the statement of this theorem. Let $\{X_i\}_{i=1}^{\infty}$ be random variables in $\mathcal{D}$ with coherent$_3$ forecasts $P(X_i) = p_i$ for $i = 1, 2, \ldots$. Assume that the forecast for X_i will be scored by a scoring rule $g_i \in \mathcal{C}$ for each i. Finally, assume that*

$$(8) \qquad P\left[\sum_{i=1}^{\infty} |X_i - p_i| \right] = V < \infty \quad and$$

$$(9) \qquad P\left[\sum_{i=1}^{\infty} g_i(X_i, p_i) \right] = W < \infty.$$

There does not exist a rival set of forecasts $\{q_i\}_{i=1}^{\infty}$ such that, for all $\omega \in \Omega$,

$$(10) \qquad \sum_{i=1}^{\infty} g_i(X_i(\omega), p_i) > \sum_{i=1}^{\infty} g_i(X_i(\omega), q_i).$$

Theorem 1 asserts conditions under which infinite sums of strictly proper scores, with coherent$_3$ forecasts $\{p_i\}_{i=1}^{\infty}$ for $\{X_i\}_{i=1}^{\infty}$, have no rival forecasts that simply dominate, let alone uniformly strictly dominate $\{p_i\}_{i=1}^{\infty}$. That is, even countably many unconditional coherent$_3$ forecasts cannot be simply dominated under the conditions of Theorem 1.

EXAMPLE 3 (Example 2 continued). Recall that $\Omega = \{\omega_i : i = 1, \ldots\}$ is a countable space. Consider the special case in which P is a purely finitely additive probability satisfying $P(\{\omega_i\}) = p_i = 0$, for all i. So, $\sum_{i=1}^{\infty} p_i = 0 < 1$, and $c = 0$ in the notation of Example 2. As before, let W_i $(i = 1, \ldots)$ be the indicator functions for the states in Ω. So $P(W_i) = p_i = 0$ and combining the losses $W_i - p_i = W_i$, for $i = 1, \ldots$ results in a uniform sure-loss of 1. But this example, with each g_i equal to Brier score times $\alpha_i > 0$, satisfies the conditions of Theorem 1, if the α_i are bounded above. That is, there are no rival forecasts $\{q_i\}_{i=1}^{\infty}$ for the $\{W_i\}_{i=1}^{\infty}$ that simply dominate the forecasts $\{p_i\}_{i=1}^{\infty}$ by weighted sum of Brier scores, let alone uniformly strictly dominating these forecasts. We can illustrate the conclusion of Theorem 1 directly in this example. The weighted sum of Brier scores for the p_i forecasts is

$$S(\omega) = \sum_{i=1}^{\infty} \alpha_i W_i(\omega)^2 \leq \sup_i \alpha_i.$$

Let $\{q_i\}_{i=1}^{\infty}$ be a rival set of forecasts with $q_i \neq p_i$ for at least one i. The corresponding weighted sum of Brier scores is

$$(11) \qquad \sum_{i=1}^{\infty} \alpha_i [W_i(\omega) - q_i]^2 = S(\omega) - 2 \sum_{i=1}^{\infty} \alpha_i q_i W_i(\omega) + \sum_{i=1}^{\infty} \alpha_i q_i^2.$$

Let $d = \sum_{i=1}^{\infty} \alpha_i q_i^2$, which must be strictly greater than 0. Define $i(\omega)$ to be the unique value of i such that $W_i(\omega) = 1$. The right-hand side of (11) can then be written as $S(\omega) - 2\alpha_{i(\omega)} q_{i(\omega)} + d$. If $d = \infty$, then the rival forecasts clearly fail to dominate the original forecasts. If $d < \infty$, then $\lim_{i \to \infty} \sqrt{\alpha_i} q_i = 0$. Because the α_i themselves are bounded, it follows that all but finitely many $\alpha_i |q_i|$ are less than $d/2$. For each ω such that $\alpha_{i(\omega)} |q_{i(\omega)}| < d/2$, we have the weighted sum of Brier scores displayed in (11) strictly greater than $S(\omega)$, hence the rival forecasts do not dominate the original forecasts.

Theorem 1, as illustrated by Example 3, shows that the modified decision problem in de Finetti's prevision game—modified to include infinite sums of betting outcomes—is not isomorphic to the modified forecasting problem under strictly proper scoring rules—modified to include infinite sums of scores. In particular, abstaining from betting, which is the alternative that uniformly dominates the losses for coherence$_1$, is not an available alternative under forecasting with strictly proper scores. In summary, the two criteria, coherence$_1$ and coherence$_3$ behave differently when probability is merely finitely additive and we try to combine countably many forecasts.

We conclude this section with an example to show why we assume that the class of scoring rules satisfies the uniform spread condition in Theorem 1.

EXAMPLE 4. This example satisfies all of the conditions of Theorem 1 except that the class of scoring rules fails the uniform spread condition. We show that the conclusion to Theorem 1 also fails. For each integer $i \geq 1$, let $\alpha_i = 2^{-i-1}$, and define

$$
f_i(v) = \begin{cases} 2, & \text{if } v \leq \alpha_i, \\ \dfrac{2}{\alpha_i}, & \text{if } v > \alpha_i. \end{cases}
$$

Let λ_i be the measure whose Radon–Nikodym derivative with respect to Lebesgue measure is f_i, and define g_i by (1) using $\lambda = \lambda_i$. The form of g_i is as follows:

$$
g_i(x, q) = \begin{cases} (x - q)^2, & \text{if } x, q \leq \alpha_i, \\ (x - \alpha_i)^2 + \dfrac{1}{\alpha_i}(q - \alpha_i)^2 + \dfrac{2}{\alpha_i}(q - \alpha_i)(\alpha_i - x), & \text{if } x \leq \alpha_i \leq q, \\ \dfrac{1}{\alpha_i}(x - \alpha_i)^2 + (q - \alpha_i)^2 + 2(\alpha_i - q)(x - \alpha_i), & \text{if } q \leq \alpha_i \leq x, \\ \dfrac{1}{\alpha_i}(x - q)^2, & \text{if } \alpha_i \leq x, q. \end{cases}
$$

These scoring rules fail the uniform spread condition because arbitrarily short intervals with both endpoints positive have arbitrarily large λ_i measure as i increases. Let $\{A_i\}_{i=1}^{\infty}$ be a partition of the real line, and let $P(\cdot)$ be a finitely additive probability such that $P(A_i) = 0$ for all i. For each integer $i \geq 1$, let $p_i = 2^{-i}$ and $q_i = 2^{-i-1}$, and define

$$
X_i(\omega) = \begin{cases} p_i, & \text{if } \omega \in A_i^C, \\ q_i - 1, & \text{if } \omega \in A_i. \end{cases}
$$

It follows that $P(X_i) = p_i$ for all i, and

$$
P\left[\sum_{i=1}^{\infty} |X_i - p_i|\right] = P\left[\sum_{i=1}^{\infty} A_i |q_i - 1 - p_i|\right] = 1,
$$

so that (8) holds. Next, compute the various scores:

$$
\begin{aligned}
& g_i\big(X_i(\omega), p_i\big) \\
& = A_i(\omega)\left\{(q_i - 1 - \alpha_i)^2 + \frac{1}{\alpha_i}(p_i - \alpha_i)^2 + \frac{2}{\alpha_i}(p_i - \alpha_i)(\alpha_i - q_i + 1)\right\} \\
& = A_i(\omega)[3 + 2^{-i-1}], \\
& g_i\big(X_i(\omega), q_i\big) \\
& = A_i^C(\omega)\frac{1}{\alpha_i}(p_i - q_i)^2 + A_i(\omega)(q_i - 1 - q_i)^2 \\
& = A_i^C(\omega)2^{-i-1} + A_i(\omega).
\end{aligned}
$$

Define $i(\omega) = i$ for that unique i such that $\omega \in A_i$. When we sum up the scores for the forecasts $\{p_i\}_{i=1}^{\infty}$, we get

$$\sum_{i=1}^{\infty} g_i(X_i(\omega), p_i) = \sum_{i=1}^{\infty} A_i(\omega)[3 + 2^{-i-1}] = 3 + 2^{-i(\omega)-1}.$$

It follows that $P(\sum_{i=1}^{\infty} g_i(X_i, p_i)) = 3$, so (9) holds. The sum of the $\{q_i\}_{i=1}^{\infty}$ scores is

$$\sum_{i=1}^{\infty} g_i(X_i(\omega), q_i) = \sum_{i=1}^{\infty} [A_i^C(\omega) 2^{-i-1} + A_i(\omega)] = 1.5 - 2^{-i(\omega)-1}.$$

Finally, compute the difference in total scores:

$$\sum_{i=1}^{\infty} g_i(X_i(\omega), p_i) - \sum_{i=1}^{\infty} g_i(X_i(\omega), q_i) = 1.5 + 2^{-i(\omega)} > 1.5,$$

hence the scores of the $\{p_i\}_{i=1}^{\infty}$ forecasts are uniformly strictly dominated by the scores of a set of rival forecasts.

4.2. *Dominance for countable sums of conditional forecasts.* Definition 2 allows mixing conditional forecasts with unconditional forecasts by setting $P(X|H) = P(X)$ whenever $H = \Omega$. De Finetti showed that, if $P(X)$, $P(X|H)$ and $P(HX)$ are all specified, a necessary condition for coherence₁ is that

(12) $$P(HX) = P(H)P(X|H),$$

so that $P(X|H)$ is the usual conditional expected value of X given H whenever $P(H) > 0$. For this reason, conditional forecasts are often called *conditional expectations*.

The concept of *conglomerability* plays a central role in our results about coherence for combining countably many conditional forecasts. Conglomerability in a partition $\pi = \{H_j : j \in J\}$ of conditional expectations $P(\cdot|H_j)$ over a class $\mathcal{D}$ of random variables X is the requirement that the unconditional expectation of each $X \in \mathcal{D}$ lies within the range of its conditional expectations given elements of π.

DEFINITION 7. Let P be a finitely additive prevision on a set $\mathcal{D}$ of random variables, and let $\pi = \{H_j : j \in J\}$ be a partition of Ω such that conditional prevision $P(\cdot|H_j)$ has been defined for all j. If, for each $X \in \mathcal{D}$,

$$\inf_{j \in J} P(X|H_j) \leq P(X) \leq \sup_{j \in J} P(X|H_j),$$

then P is *conglomerable in the partition π with respect to $\mathcal{D}$*. Otherwise, P is *nonconglomerable* in π with respect to $\mathcal{D}$.

If a decision maker's coherent$_1$ or coherent$_3$ forecasts fail conglomerability in a partition π, Theorem 2 below shows there exist countably many conditional forecasts that are uniformly strictly dominated.

On the other hand, if the decision maker's previsions for random variables satisfy a condition (see Definition 8) similar to being conglomerable in π, Theorem 3 below establishes that no countable set of forecasts, conditional on elements of π, can be uniformly strictly dominated. What we mean by "similar" is explained in Section 4.3 below.

THEOREM 2. *Let P be a finitely additive prevision, and let $\mathcal{D}$ be a set of random variables. Let $\pi = \{H_j\}_{j=1}^{\infty}$ be a denumerable partition and let $P(\cdot|H_j)$ be the corresponding conditional previsions associated with P. Let $\mathcal{C}$ be a collection of strictly proper scoring rules of the form (1) that satisfies the uniform similarity condition. Assume that the conditional previsions $P(\cdot|H_j)$ are nonconglomerable in π with respect to $\mathcal{D}$. Then there exists a random variable $X \in \mathcal{D}$ with $p_X = P(X)$ and $p_j = P(X|H_j)$ for all j such that*
(2.1) the countable sum

$$\alpha_0(X - p_X) + \sum_{j=1}^{\infty} \alpha_j H_j(X - p_j),$$

of individually fair options is uniformly strictly dominated by abstaining, and
(2.2) if the forecast for X is scored by $g_0 \in \mathcal{C}$ and the conditional forecast for X given H_j is scored by $g_j \in \mathcal{C}$ for $j = 1, 2, \ldots$, then the sum of the scores,

$$g_0\big(X(\omega), p_X\big) + \sum_{j=1}^{\infty} H_j(\omega)g_j\big(X(\omega), p_j\big),$$

is uniformly strictly dominated by the sum of scores from a rival set of forecasts.

We illustrate Theorem 2 with an example of nonconglomerability due to Dubins (1975). This example is illuminating as the conditional probabilities do not involve conditioning on null events.

EXAMPLE 5. Let $\Omega = \{\omega_{ij} : i = 1, 2; j = 1, \ldots\}$. Let $F = \{\omega_{2j}, j = 1, \ldots\}$ and let $H_j = \{\omega_{1j}, \omega_{2j}\}$. Define a merely finitely additive probability P so that $P(\{\omega_{1j}\}) = 0$, $P(\{\omega_{2j}\}) = 2^{-(j+1)}$ for $j = 1, \ldots$, and let $P(F) = p_F = 1/2$. Note that $P(H_j) = 2^{-(j+1)} > 0$, so $P(F|H_j) = 1 = p_j$ is well defined by the multiplication rule for conditional probability. Evidently, the conditional probabilities $\{P(F|H_j)\}_{j=1}^{\infty}$ are nonconglomerable in π since $P(F) = 1/2$ whereas $P(F|H_j) = 1$ for all j.

For (2.1), Consider the fair options $\alpha_j H_j (F - p_j)$ for $j = 1, \ldots$ and $\alpha_F (F - p_F)$. Choose $\alpha_j = 1$ and $\alpha_F = -1$. Then

$$\left[-(F(\omega) - p_F) + \sum_{j=1}^{\infty} H_j(\omega) \big[F(\omega) - p_j \big] \right]$$

$$= \begin{cases} 0.5 - 1.0 = -0.5, & \text{if } \omega \notin F, \\ -0.5 + 0.0 = -0.5, & \text{if } \omega \in F. \end{cases}$$

Hence, these infinitely many individually fair options are not collectively fair when taken together. Their sum is uniformly strictly dominated by 0 in Ω, corresponding to the option to abstain from betting.

Regarding (2.2), unlike the situation with Theorem 1 involving countably many unconditional forecasts, the sum of Brier scores from these conditional forecasts are uniformly strictly dominated. In particular, the sum of Brier scores for these forecasts is

$$\big(F(\omega) - p_F \big)^2 + \sum_{j=1}^{\infty} H_j(\omega) \big[F(\omega) - p_j \big]^2$$

$$= \begin{cases} 0.25 + 1.00 = 1.25, & \text{if } \omega \notin F, \\ 0.25 + 0.00 = 0.25, & \text{if } \omega \in F. \end{cases}$$

Consider the rival forecasts $Q(F|H_j) = 0.75 = q_j$ and $Q(F) = 0.75 = q_F$. These correspond to the countably additive probability $Q(\{\omega_{1j}\}) = 0.25 \times 2^{-j}$ and $Q(\{\omega_{2j}\}) = 0.75 \times 2^{-j}$ for $j = 1, \ldots$. Then the combined Brier score from these countably many rival forecasts is

$$\big(F(\omega) - q_F \big)^2 + \sum_{j=1}^{\infty} H_j(\omega) \big[F(\omega) - q_j \big]^2$$

$$= \begin{cases} 9/16 + 9/16 = 1.125, & \text{if } \omega \notin F, \\ 1/16 + 1/16 = 0.125, & \text{if } \omega \in F, \end{cases}$$

which is 0.125 less than the sum of the Brier scores of the original forecasts.

We offer one more example to show why we assume that the class of scoring rules satisfies the uniform similarity condition in Theorem 2.

EXAMPLE 6 (Example 5 continued). Recall that we have a partition $\pi = \{H_j\}_{j=1}^{\infty}$ and an event F with $p_F = P(F) = 0.5$ and $p_j = P(F|H_j) = 1$ for all j. Let the unconditional forecast for F be scored by Brier score, and let the conditional forecast for F given H_j be scored by 2^{-j-1} times Brier score. These scoring rules fail the uniform similarity condition. We show that the conclusion to Theorem 2 fails. Specifically, we show that there is no rival set of forecasts q_F for F

and q_j for H_j ($j = 1, 2, \ldots$) whose sum of scores uniformly strictly dominates the original forecasts.

The total of the scores for the original forecasts is

$$(13) \qquad \frac{1}{4} + \sum_{j=1}^{\infty} H_j(\omega) 2^{-j-1} [1 - F(\omega)]^2.$$

Consider an arbitrary rival set of forecasts with q_F for F and q_j for F conditional on H_j. The sum of the scores for the rival forecasts is

$$(14) \qquad [q_F - F(\omega)]^2 + \sum_{j=1}^{\infty} H_j(\omega) 2^{-j-1} [q_j - F(\omega)]^2.$$

Let $i(\omega) = j$ when $\omega \in H_j$. Then the difference (13) minus (14) is

$$(15) \qquad \tfrac{1}{4} - [q_F - F(\omega)]^2 + 2^{-i(\omega)-1} ([1 - F(\omega)]^2 - [q_j - F(\omega)]^2).$$

If $q_F = 0.5$, then (15) becomes

$$(16) \qquad 2^{-i(\omega)} (1 - q_{i(\omega)}) \left[\frac{1 + q_{i(\omega)}}{2} - F(\omega) \right].$$

If there exists ω such that $q_{i(\omega)} \geq 1$, (16) is nonpositive, and the rival forecasts do not strictly dominate. If all $q_{i(\omega)} < 1$, (16) is negative for all $\omega \in F$, and there is no dominance. If $q_F \neq 0.5$, then (15) is at most

$$(17) \qquad \tfrac{1}{4} - [q_F - F(\omega)]^2 + 2^{-i(\omega)}.$$

No matter what $q_F \neq 0.5$ we pick, either $(q_F - 1)^2$ or $(q_F - 0)^2$ is greater than $1/4$. Let $\delta = 1/4 - \max\{[q_F - 1]^2, q_F^2\}$. For $j > -\log_2(\delta)$, (17) is negative either for all $\omega \in F \cap H_j$ or all $\omega \in F^C \cap H_j$. So, there is no dominance.

Last, we establish conditions under which combining strictly proper scores from countably many conditional forecasts given elements of a partition, or combining the losses from countably many fair options based on those forecasts, does not result in a uniform sure loss. A definition is useful first.

DEFINITION 8. Let P be a finitely additive prevision on a set $\mathcal{D}$ of random variables, and let $\pi = \{H_j : j \in J\}$ be a partition of Ω such that conditional prevision $P(\cdot | H_j)$ has been defined for all j. For each random variable $X \in \mathcal{D}$, we let $P(X | \pi)$ denote the random variable Y defined by $Y(\omega) = P(X | H_j)$ for all $\omega \in H_j$ and all j. We say that P satisfies the law of total previsions in π with respect to $\mathcal{D}$ provided that for each random variable $X \in \mathcal{D}$, $P(X) = P[P(X | \pi)]$.

THEOREM 3. *Let P be a finitely additive prevision, and let $\mathcal{D}$ be a set of random variables such that P satisfies the law of total previsions in $\pi = \{H_j\}_{j=1}^{\infty}$ with respect to $\mathcal{D}$. Let $X \in \mathcal{D}$ be a random variable with finite prevision $p_X = P(X)$ and finite conditional prevision $p_j = P(X|H_j)$ given each H_j. Assume that $P(\cdot|H_j)$ is a finitely additive expectation (in the sense of Definition 10) that satisfies $P(X|H_j) = P(H_j X|H_j)$ for every j. Let $\mathcal{C}$ be a collection of strictly proper scoring rules.*

(3.1) *Let $\{\alpha_j\}_{j=0}^{\infty}$ be real numbers. The sum of losses*

$$(18) \qquad \alpha_0\big(X(\omega) - p_X\big) + \sum_{j=1}^{\infty} \alpha_j H_j(\omega)\big[X(\omega) - p_j\big],$$

is not uniformly strictly dominated by abstaining.

(3.2) *Let $g_0, g_1, \ldots$ be elements of $\mathcal{C}$. There is no rival set of forecasts that uniformly strictly dominates the sum of scores*

$$(19) \qquad g_0\big(X(\omega), p_X\big) + \sum_{j=1}^{\infty} H_j(\omega) g_j\big(X(\omega), p_j\big).$$

4.3. *Conglomerability, disintegrability and the law of total previsions.* We claimed earlier that the law of total previsions in a partition π is similar to conglomerability in π. The claim begins with a result of Dubins (1975). Dubins defines *conglomerability in partition π* of a finitely additive prevision P by the requirement that, for all bounded random variables X,

$$\text{if } \forall H \in \pi \, P(X|H) \geq 0, \text{ then } P(X) \geq 0.$$

Dubins' definition of conglomerability in π is equivalent to Definition 7 with respect to the set of all bounded random variables. However, for a set $\mathcal{D}$ that includes unbounded random variables and/or does not include all bounded random variables, the two definitions are not equivalent without further assumptions. Definition 7 is based on the definition given by de Finetti [(1974), Section 4.7], which generalizes to unbounded random variables more easily.

Dubins (1975) also defines *disintegrability of P in partition π* by the requirement that, for every bounded random variable X,

$$P(X) = \int P(X|h) \, dP(h),$$

where the finitely additive integral is as developed by Dunford and Schwartz [(1958), Chapter III]. Moreover, he establishes that conglomerability and disintegrability in π are equivalent for the class of bounded random variables.

The law of total previsions in Definition 8, with respect to the set of all bounded random variables, is equivalent to disintegrability in Dubins' sense, but not necessarily for sets that either include some unbounded random variables or fail to

include some bounded random variables. In addition, not all real-valued coherent$_1$ previsions admit an integral representation in the sense of Dunford and Schwartz for sets that include unbounded random variables. For discussion of the problem and related issues, see Berti, Regazzini and Rigo (2001); Berti and Rigo (1992, 2000, 2002); Schervish, Seidenfeld, and Kadane (2008b) and Seidenfeld, Schervish and Kadane (2009). As described in Appendix A, we use a definition of finitely additive integral that is a natural extension of coherent$_1$ prevision. In this way, the law of total previsions extends Dubins' definition of disintegrability from bounded to unbounded random variables without introducing the technical details of Dunford and Schwartz. Finally, Theorem 1 of Schervish, Seidenfeld and Kadane (2008b) gives conditions under which conglomerability (Definition 7) is equivalent to the law of total previsions. The following is a translation of that result into the notation and terminology of the present paper.

THEOREM 4. *Let P be a finitely additive prevision on a set $\mathcal{D}$ of random variables. Let $\pi = \{H_j\}_{j=1}^{\infty}$ be a denumerable partition and let $P(\cdot|H_j)$ be the corresponding conditional previsions associated with P. Assume that, for all j, $P(\cdot|H_j)$ is a finitely additive expectation on $\mathcal{D}$. Also assume that, for all $X \in \mathcal{D}$:*

- *$P(X)$ is finite,*
- *$P(X|H_j)$ is finite for all j,*
- *$H_j X \in \mathcal{D}$ for all j,*
- *$P(H_j X|H_j) = P(X|H_j)$ for all j, and*
- *$X - Y \in \mathcal{D}$, where Y is defined (in terms of X) in Definition 8.*

Then P is conglomerable in π with respect to $\mathcal{D}$ if and only if P satisfies the law of total previsions in π with respect to $\mathcal{D}$.

Under the conditions of Theorem 4, Theorems 2 and 3 show that, when the conditioning events form a countable partition π, coherence$_1$ and coherence$_3$ behave the same when extended to include, respectively, the countable sum of individually fair options, and the total of strictly proper scores from the forecasts. If and only if these coherent quantities are based on conditional expectations that are conglomerable in π, then no failures of the Dominance Principle result by combining infinitely many of them.

Schervish, Seidenfeld and Kadane (1984) show that each merely finitely additive probability fails to be conglomerable in some countable partition. But each countably additive probability has expectations that are conglomerable in each countable partition. Thus, the conjunction of Theorems 1, 2 and 3 identifies where the debate whether personal probability may be merely finitely additive runs up against the debate whether to extend either coherence criterion in order to apply it with countable combinations of quantities. We arrive at the following conclusions:

- Unless unconditional coherent$_1$ forecasts arise from a countably additive probability, combining countably many unconditional coherent$_1$ forecasts into a single option may be dominated by abstaining.
- However, under the conditions of Theorem 1, strictly proper scoring rules are not similarly affected. The scores from countably many coherent$_3$ unconditional forecasts may be summed together without leading to a violation of the Dominance Principle.
- Unless conditional forecasts arise from a set of conglomerable conditional probabilities, the Dominance Principle does not allow combining countably many of these quantities into a single option. Hence, only countably additive conditional probabilities satisfy the Dominance Principle when an arbitrary countable set of conditional quantities are summed together.

5. Incentive compatible elicitation of infinitely many forecasts using strictly proper scoring rules. Scoring an agent based on the values of the fair gambles constructed from coherent$_1$ forecasts, is not proper. Because of the presence of the opponent in the game, who gets to choose whether to buy or to sell the random variable X at the decision maker's announced price, the decision maker faces a strategic choice of pricing. For example, if the decision maker suspects that the opponent's fair price, $Q(X)$, is greater than his own, $P(X)$, then it pays to inflate the announced price and to offer the opponent, for example, $R(X) = [P(X) + Q(X)]/2$, rather than offering $P(X)$. Thus, the forecast-game as de Finetti defined it for coherence$_1$ is not incentive compatible for eliciting the decision maker's fair prices.

With a finite set of forecasts and a strictly proper scoring rule for each one, using the finite sum of the scores as the score for the finite set preserves strict propriety. That is, with the sum of strictly proper scores as the score for the finite set, a coherent forecaster minimizes the expected sum of scores by minimizing each one, and this solution is unique.

Here, we report what happens to the propriety of strictly proper scores in each of the three settings of the three theorems presented in Section 4. That is, we answer the question whether or not, in each of these three settings, the coherent forecaster minimizes expected score for the infinite sum of strictly proper scores by announcing her/his coherent forecast for each of the infinitely many variables. These findings are corollaries to the respective theorems.

COROLLARY 1. *Under the assumptions of Theorem 1, the infinite sum of scores applied to the infinite set of forecasts $\{p_i\}_{i=1}^{\infty}$ is a strictly proper scoring rule.*

COROLLARY 2. *Under the assumptions used for (2.2) of Theorem 2, namely when the conditional probabilities $P(F|H_j) = p_j$ are nonconglomerable in π, then the infinite sum of strictly proper scores applied to the infinite set of conditional forecasts $\{p_j\}_{j=1}^{\infty}$ is not proper.*

COROLLARY 3. *Under the assumptions used to establish* (3.2) *of Theorem* 3, *namely that* P *satisfies the law of total previsions in* π, *the infinite sum of strictly proper scores applied to the infinite set of conditional forecasts* $\{p_j\}_{j=1}^{\infty}$ *is a proper scoring rule.*

Thus, these results about the propriety of infinite sums of strictly proper scores parallel the respective results about extending coherence₃ to allow infinite sums of scores.

6. Summary. We study how two different coherence criteria behave with respect to a Dominance Principle when countable collections of random variables are included. Theorem 1 shows that, in contrast with fair prices for coherence₁, when strictly proper scores from infinitely many unconditional forecasts are summed together there are no new failures of the Dominance Principle for coherence₃. That is, if an infinite set of probabilistic forecasts $\{p_i\}_{i=1}^{\infty}$ are even simply dominated by some rival forecast scheme $\{q_i\}_{i=1}^{\infty}$ in total score, then the $\{p_i\}_{i=1}^{\infty}$ are not coherent₃, that is, some finite subset of them is uniformly strictly dominated in total score. However, because each merely finitely additive probability fails to be conglomerable in some denumerable partition, in the light of Theorem 2, neither of the two coherence criteria discussed here may be relaxed in order to apply the Dominance Principle with infinite combinations of conditional options. Merely finitely additive probabilities then would become *incoherent*.

Specifically, the conjunction of Theorems 1–4 shows that it matters only in cases that involve nonconglomerability whether incoherence₃ is established using scores from a finite rather than from an infinite combination of forecasts. In that one respect, we think coherence₃ constitutes an improved version of the concept of coherence. Coherence₁ applied to a merely finitely additive probability leads to failures of the Dominance Principle both with infinite combinations of unconditional and infinite combinations of nonconglomerable conditional probabilities. Coherence₃ leads to failures of the Dominance Principle only with infinite combinations of nonconglomerable conditional probabilities.

A referee suggested that de Finetti might have been working with a different Dominance Principle, here denoted Dominance*.

> *Dominance**: Let O_i and O_j be two options in $\mathcal{O}$. If O_i uniformly (strictly) dominances O_j and there exists an option O_k in $\mathcal{O}$ that is not itself dominated by some O_t in $\mathcal{O}$, then O_j is an inadmissible choice from $\mathcal{O}$.*

Dominance* requires that some option from $\mathcal{O}$ is undominated if dominance signals inadmissibility. With respect to the decision problems considered in this paper, each of our results formulated with respect to the Dominance Principle obtains also with Dominance*. Because Dominance* implies Dominance as we have defined it, the only result that needs to be checked is Theorem 2. In that case, so long as $\mathcal{O}$ contains options that correspond to a probability that satisfies the law of total previsions in π (as will all countably additive probabilities) then Theorem 3

says that such options will be undominated. So, we could replace Dominance by Dominance* in the results of this paper.

APPENDIX A: FINITELY ADDITIVE EXPECTATIONS

This appendix gives the definitions of infinite prevision and finitely additive expectation along with brief motivation for these definitions. Details are given in the supplemental article [Schervish, Seidenfeld and Kadane (2014)].

A.1. Infinite previsions.
Our theorems assume that various random variables have finite previsions. In the proof of Theorem 1, the possibility arises that some other random variable has infinite prevision. Definition 2 makes no sense if infinite previsions are possible. Fortunately, we can extend the concept of coherent$_1$ (conditional) prevision to handle infinite values, which correspond to expressing a willingness either to buy or to sell a gamble, but not both.

DEFINITION 9. Let $\{P(X_i|B_i) : i \in I\}$ be a collection of conditional previsions. The previsions are *coherent$_1$* if, for every finite n, every $\{i_1, \ldots, i_n\} \subseteq I$, all real $\alpha_1, \ldots, \alpha_n$ such that $\alpha_j \leq 0$ for all j with $P(X_{i_j}|B_{i_j}) = \infty$ and $\alpha_j \geq 0$ for all j with $P(X_{i_j}|B_{i_j}) = -\infty$, and all real $c_1, \ldots, c_n$ such that $c_j = P(X_{i_j}|B_{i_j})$ for each j such that $P(X_{i_j}|B_{i_j})$ is finite, we have

$$(20) \qquad \inf_{\omega \in \Omega} \sum_{j=1}^{n} \alpha_j B_{i_j}(\omega)\big[X_{i_j}(\omega) - c_j\big] \leq 0.$$

That is, no linear combination of gambles may be uniformly strictly dominated by the alternative option of abstaining.

Notice the restrictions on the signs of coefficients in Definition 9, namely that for each infinite prevision, α_j has the opposite sign as the prevision. These restrictions express the meaning of infinite previsions as being *one-sided* in the sense that they merely specify that all real numbers are either acceptable buy prices (for ∞ previsions) *or* acceptable sell prices (for $-\infty$ previsions) but not fair prices for both transactions. Crisma, Gigante and Millossovich (1997) and Crisma and Gigante (2001) give alternate definitions of coherence for infinite previsions and conditional previsions. But their definition does not make clear the connection to gambling. However, the definition of Crisma, Gigante and Millossovich (1997) and Definition 9 are equivalent for unconditional previsions, as shown in the supplemental article [Schervish, Seidenfeld and Kadane (2014)].

A.2. Prevision and expectation.
Throughout this paper, an expectation with respect to a finitely additive probability will be defined as a special type of linear functional on a space of random variables. [See Heath and Sudderth (1978) for the case of bounded random variables.] Infinite previsions are allowed in the sense of Section A.1.

DEFINITION 10.   Let $\mathcal{L}$ be a linear space of real-valued functions defined on Ω that contains all constant functions, and let L be an extended-real-valued functional defined on $\mathcal{L}$. If $(X, Y \in \mathcal{L}$ and $X \leq Y)$ implies $L(X) \leq L(Y)$, we say that L is *nonnegative*. We call L an *extended-linear functional* on $\mathcal{L}$, if, for all real α, β and all $X, Y \in \mathcal{L}$,

$$(21) \qquad L(\alpha X + \beta Y) = \alpha L(X) + \beta L(Y),$$

whenever the arithmetic on the right-hand side of (21) is well defined (i.e., not $\infty - \infty$) and where $0 \times \pm\infty = 0$ in (21). A nonnegative extended-linear functional is called a *finitely additive Daniell integral*. [See Schervish, Seidenfeld and Kadane (2008a).] If $L(1) = 1$, we say that L is *normalized*. A normalized finitely additive Daniell integral is called a *finitely additive expectation*.

Note that, if $\infty - \infty$ appears on the right-hand side of (21), $L(\alpha X + \beta Y)$ still has a value, but the value cannot be determined from (21). Finitely additive expectations are essentially equivalent to coherent$_1$ previsions, as we prove in the supplemental article. Finitely additive expectations also behave like integrals in many ways, as we explain in more detail in the supplemental article. In particular, when the finitely additive expectation defined here is restricted to bounded functions, it is the same as the definition of integral developed by Dunford and Schwartz (1958), and it is the same as the integral used by Dubins (1975) in his results about disintegrability. Hence, Definition 10 is an extension of the definition of integral from sets of bounded functions to arbitrary linear spaces of functions.

APPENDIX B: PROOFS OF RESULTS

B.1. Proof of Lemma 1.   Let g be of the form (1). Let P be such that $p = P(X)$ is finite, and let q_0 be such that $P[g(X, q_0)]$ is finite. If $q \neq p$, then

$$(22) \qquad P[g(X, q) - g(X, p)] = \lambda((q, p))[p - r(q, p, \lambda)],$$

according to (5). Because Lebesgue measure is absolutely continuous with respect to λ, neither $\lambda((q, p))$ nor $p - r(q, p, \lambda)$ equals 0 and they have the same sign. It follows that (22) is strictly positive. Since p is finite, (22) is finite with $q = q_0$, so that $P[g(X, p)]$ is also finite and so $q = p$ provides the unique minimum value of $P[g(X, q)]$.

B.2. Proof of Lemma 2.   Since $r(b, a, \lambda) = r(a, b, \lambda)$, it suffices to assume that $a < b$. Let $\varepsilon > 0$.

(i) If $\lambda_i((a, b)) \geq \varepsilon$ and $b_0 < b$ is such that $\lambda_i((a, b_0)) = \varepsilon$, then the probability obtained by normalizing λ_i on the interval (a, b) stochastically dominates the probability obtained by normalizing λ_i on the interval (a, b_0). Hence, $r(a.b, \lambda_i) \geq r(a, b_0, \lambda_i)$. So, it suffices to find a δ that implies $r(a, b, \lambda_i) - a \geq \delta$ for all $i \in I$ and all $a < b$ such that $\lambda_i((a, b)) = \varepsilon$. For the remainder of the proof,

let $a < b$ with $\lambda_i((a, b)) = \varepsilon$, and let Q be the probability obtained by normalizing λ_i on (a, b). Let λ_0 be U times Lebesgue measure. Then $\lambda_0((a, a + \varepsilon/U)) = \varepsilon$, and $r(a, a + \varepsilon/U, \lambda_0) = a + \varepsilon/(2U)$. Because $f_i \leq U$, it follows that Q stochastically dominates the probability obtained by normalizing λ_0 on $(a, a + \varepsilon/U)$, hence $r(a, b, \lambda_i) \geq a + \varepsilon/(2U)$, and $r(a, b, \lambda_i) - a \geq \varepsilon/(2U)$. The proof $b - r(a, b, \lambda_i) \geq \varepsilon/(2U)$ is similar, so δ_ε can be taken equal to $\varepsilon/(2U)$.

(ii) Let $i \in I$, and assume that $\lambda_i((a, b)) \geq \varepsilon$. Since $f_j(v) > L_i f_i(v)$ for all v, we have $\lambda_j((a, b)) \geq L_i \varepsilon$, so $\gamma_{i,\varepsilon}$ can be taken to be $L_i \varepsilon$.

B.3. Proofs of Theorem 1 and Corollary 1. Because a larger random variable has a larger prevision than a smaller random variable, a necessary condition for (10) is that

$$(23) \qquad Z = P\left[\sum_{i=1}^{\infty} g_i(X_i, q_i)\right] \leq P\left[\sum_{i=1}^{\infty} g_i(X_i, p_i)\right] < \infty.$$

Hence, we will assume that $Z < \infty$ from now on. Also, it is necessary for (10) that $q_i \neq p_i$ for at least one i, so we will assume this also.

In light of (5), we can write, for each finite $k > 0$,

$$\infty > Z - W = P\left[\sum_{i=1}^{\infty} g_i(X_i, q_i) - \sum_{i=1}^{\infty} g_i(X_i, p_i)\right]$$

$$= \sum_{i=1}^{k} \lambda_i((q_i, p_i))[p_i - r_i] + P\left[\sum_{i=k+1}^{\infty} g_i(X_i, q_i) - \sum_{i=k+1}^{\infty} g_i(X_i, p_i)\right]$$

$$\geq \sum_{i=1}^{k} \lambda_i((q_i, p_i))[p_i - r_i] - W,$$

where the inequality follows because g_i is nonnegative for each i and where $r_i = r(q_i, p_i, \lambda_i)$ from (4). Since $Z - W$ does not depend on k, it follows that $\sum_{i=1}^{\infty} \lambda_i((q_i, p_i))[p_i - r_i]$ is finite.

Because of (9) and (23), the two series $\sum_{i=1}^{\infty} g_i(X_i(\omega), q_i)$ and $\sum_{i=1}^{\infty} g_i(X_i(\omega), p_i)$ are simultaneously finite with probability 1. Let B be the event that at least one of the two series is finite. On B^C, both series sum to ∞, hence (10) fails unless $B^C = \varnothing$. Hence, we can assume that $B = \Omega$ for the rest of the proof. It now follows that, for all ω,

$$(24) \qquad \sum_{i=1}^{\infty} g_i(X_i(\omega), q_i) - \sum_{i=1}^{\infty} g_i(X_i(\omega), p_i)$$

$$= \sum_{i=1}^{\infty} [g_i(X_i(\omega), q_i) - g_i(X_i(\omega), p_i)].$$

We complete the proof by showing that

$$(25) \qquad P\left[\sum_{i=1}^{\infty} g_i(X_i, q_i) - \sum_{i=1}^{\infty} g_i(X_i, p_i)\right] > 0.$$

Because a nonpositive random variable has nonpositive forecast, (25) implies that (10) cannot hold for all ω. In light of (24), it suffices to show that

$$(26) \qquad P\left(\sum_{i=1}^{\infty}[g_i(X_i, q_i) - g_i(X_i, p_i)]\right) > 0.$$

For each k,

$$(27) \qquad P\left(\sum_{i=1}^{k}[g_i(X_i, q_i) - g_i(X_i, p_i)]\right) = \sum_{i=1}^{k} \lambda_i((q_i, p_i))(p_i - r_i) \geq 0.$$

Next, in light of (5) and (27), write

$$(28)$$
$$P\left(\sum_{i=1}^{\infty}[g_i(X_i, q_i) - g_i(X_i, p_i)]\right)$$
$$= \sum_{i=1}^{k} \lambda_i((q_i, p_i))(p_i - r_i) + P\left[\sum_{i=k+1}^{\infty} \lambda_i((q_i, p_i))(X_i - r_i)\right].$$

Since the left-hand side of (28) does not depend on k and the first sum on the right side is nondecreasing in k, it follows that the second sum on the right-hand side is nonincreasing in k, and hence, has a limit. Let $T = \sum_{i=1}^{\infty} \lambda_i((q_i, p_i))(p_i - r_i)$, which is finite and strictly positive (because $q_i \neq p_i$ for at least one i). Then, the right-hand side of (28) becomes

$$(29) \qquad T + \lim_{k \to \infty} P\left[\sum_{i=k+1}^{\infty} \lambda_i((q_i, p_i))(X_i - p_i)\right].$$

The proof will be complete if we can show that the limit in (29) is 0.

First, we show that $\lim_{i \to \infty} \lambda_i((q_i, p_i)) = 0$. If $\limsup_{i \to \infty} |\lambda_i((q_i, p_i))| = \ell > 0$, then there must exist a subsequence $\{i_j\}_{j=1}^{\infty}$ with $|\lambda_{i_j}((q_{i_j}, p_{i_j}))| > \ell/2$ for all j. For such a subsequence, the uniform spread condition implies that there is $\delta_{\ell/2} > 0$ such that $|p_{i_j} - r_{i_j}| \geq \delta_{\ell/2}$. This would make $T = \infty$, a contradiction.

It now follows that

$$\left| P\left[\sum_{i=k+1}^{\infty} \lambda_i((q_i, p_i))(X_i - p_i)\right] \right| \leq \max_{i \geq k+1} |\lambda_i((q_i, p_i))| P\left(\sum_{i=1}^{\infty} |X_i - p_i|\right)$$
$$= V \max_{i \geq k+1} |\lambda_i((q_i, p_i))|,$$

which can be made arbitrarily small by increasing k, and (26) follows.

Corollary 1 is equivalent to equation (25), which is established in the proof of Theorem 1.

B.4. Proofs of Theorem 2 and Corollary 2. Let $\pi = \{H_j\}_{j=1}^{\infty}$ be a denumerable partition. Nononglomerability means that there exists a random variable X such that either

$$\inf_j P(X|H_j) - P(X) > 0 \quad \text{or}$$

$$\sup_j P(X|H_j) - P(X) < 0.$$

Clearly, if X satisfies one of the above inequalities, $-X$ satisfies the other, hence we will assume that the first inequality holds. Specifically, let $p_X = P(X)$ and $p_j = P(X|H_j)$ for all j, and assume that

$$\varepsilon = \inf_j p_j - p_X > 0.$$

Also, for each $\omega \in \Omega$, let $i(\omega)$ be the unique integer such that $\omega \in H_{i(\omega)}$. Hence $H_j(\omega) = 1$ if and only if $j = i(\omega)$.

(2.1) Consider the following sum of individually fair options: $X(\omega) - p_X$ and the countably many options $-H_j(\omega)[X(\omega) - p_j]$ for $j = 1, 2, \ldots$. Then, for each ω,

$$X(\omega) - p_X + \sum_{j=1}^{\infty} -H_j(\omega)\big[X(\omega) - p_j\big]$$

$$= X(\omega) - p_X - X(\omega) + p_{i(\omega)} = -p_X + p_{i(\omega)} \geq \varepsilon.$$

Thus, the countable sum of the conditional forecasts for X given H_j, combined with the forecast for X results in a loss that is uniformly strictly dominated by 0.

(2.2) For an arbitrary set of forecasts s_X for X and s_j for X given H_j (for $j = 1, \ldots$), the sum of the scores in state ω equals

$$g_0\big(X(\omega), s_X\big) + \sum_{j=1}^{\infty} H_j(\omega) g_j\big(X(\omega), s_j\big)$$

(30) $$= g_0\big(X(\omega), s_X\big) + g_{i(\omega)}\big(X(\omega), s_{i(\omega)}\big)$$

$$= \int_{s_X}^{X(\omega)} \big[X(\omega) - v\big] d\lambda_0(v) + \int_{s_{i(\omega)}}^{X(\omega)} \big[X(\omega) - v\big] d\lambda_{i(\omega)}(v).$$

We can substitute the original forecasts $s_X = p_X$ and $s_j = p_j$, $j = 1, \ldots$ into (30) to obtain the total score for each $\omega \in \Omega$. We can also identify dominating rival forecasts q_X and q_j, $j = 1, \ldots$, so that (30) is uniformly larger, for each state $\omega \in \Omega$ with $s_X = p_X$ and $s_j = p_j$ than with $s_X = q_X$ and $s_j = q_j$.

Let $w_0 = \lambda_0((p_X, p_X + \varepsilon))/2$, and let $w_1 = \gamma_{w_0}$, where γ_{w_0} is from part (ii) of Lemma 2. Let q' be such that $\lambda_0((q', p_X + \varepsilon)) = w_0$. This makes $\lambda_j((q', p_X + \varepsilon)) \geq w_1$ for all j. For each j, $p_j \geq p_X + \varepsilon$, so that $\lambda_j((q', p_j)) \geq w_1$. Let $w_2 =$

$0.9 \min\{w_0, w_1\}$, and let q_j be such that $\lambda_j((q_j, p_j)) = w_2$ for all j. This makes $q_j > q'$ for all j. Let q_X be such that $\lambda_0((p_X, q_X)) = w_2$. This makes $q_X < q'$.

We now form the difference between the scores for the original forecasts and the rival forecasts. Subtracting (30) with $s = q_X$ and $s_j = q_j$ (for all j) from (30) with $s = p_X$ and $s_j = p_j$ (for all j) yields

$$(31) \qquad \int_{p_X}^{q_X} \left[X(\omega) - v\right] d\lambda_0(v) - \int_{q_{i(\omega)}}^{p_{i(\omega)}} \left[X(\omega) - v\right] d\lambda_{i(\omega)}(v).$$

We need to find a positive number δ such that (31) is strictly greater than δ for all ω. The difference in (31) is greater than

$$\left[X(\omega) - q_X\right]\lambda_0((p_X, q_X)) - \left[X(\omega) - q_{i(\omega)}\right]\lambda_{i(\omega)}((q_{i(\omega)}, p_{i(\omega)}))$$
$$= w_2(q_{i(\omega)} - q_X) > w_2(q' - q_X) > 0.$$

So, we set $\delta = w_2(q' - q_X) > 0$, which completes the proof.

Corollary 2 is immediate from (2.2) of Theorem 2, as the existence of the rival set of dominating forecasts, $\{q_j\}_{j=1}^{\infty}$, establishes that the forecaster does not minimize the infinite sum of expected scores by giving the forecast p_X and the conditional forecasts $\{p_j\}_{j=1}^{\infty}$.

B.5. Proofs of Theorem 3 and Corollary 3. (3.1) In order to show that (18) cannot be uniformly strictly positive, it is sufficient to show

$$(32) \qquad P\left[\alpha_0(X - p_X) + \sum_{j=1}^{\infty} \alpha_j H_j(X - p_j)\right] = 0.$$

Of course,

$$P\left[\alpha_0(X - p_X) + \sum_{j=1}^{\infty} \alpha_j H_j(X - p_j)\right] = P[\alpha_0(X - p_X)] + P\left[\sum_{j=1}^{\infty} \alpha_j H_j(X - p_j)\right].$$

Trivially,

$$(33) \qquad P[\alpha_0(X - p_X)] = 0.$$

Since P satisfies the law of total previsions in π,

$$P\left[\sum_{j=1}^{\infty} \alpha_j H_j(X - p_j)\right] = P\left[P\left[\sum_{j=1}^{\infty} \alpha_j H_j(X - p_i)\Big|\pi\right]\right].$$

For each i, $\sum_{j \neq i} \alpha_j H_j(\omega)[X(\omega) - p_j] = 0$ for all $\omega \in H_i$. It follows that, for every i,

$$P\left[\sum_{j=1}^{\infty} \alpha_j H_j(X - p_j)\Big|H_i\right] = P[\alpha_i H_i(X - p_i)|H_i],$$

and trivially, $P[\alpha_i H_i(X - p_i)|H_i] = 0$.

Thus,

$$P\left[\sum_{j=1}^{\infty} \alpha_j H_j(X - p_i)\middle|\pi\right] = 0$$

for all ω, and it follows by the law of total previsions that

(34) $$P\left[\sum_{j=1}^{\infty} \alpha_j H_j(X - p_j)\right] = 0.$$

Equations (33) and (34) establish (32).

(3.2) We must establish that there is no rival set of forecasts q_X, and $\{q_j\}_{j=1}^{\infty}$ whose total score uniformly dominates (19). That is, there is no rival set of forecasts such that for some $\varepsilon > 0$ and every ω,

$$g_0\big(X(\omega), p_X\big) + \sum_{j=1}^{\infty} H_j(\omega)g_j\big(X(\omega), p_j\big)$$

$$\geq g_0\big(X(\omega), q_X\big) + \sum_{j=1}^{\infty} H_j(\omega)g_j\big(X(\omega), q_j\big) + \varepsilon.$$

It is sufficient to show that

$$P\left\{g_0(X, q_X) + \sum_{j=1}^{\infty} H_j g_j(X, q_j) - \left[g_0(X, p_X) + \sum_{j=1}^{\infty} H_j g_j(X, p_j)\right]\right\}$$

(35)
$$\geq 0.$$

Write the left-hand side of (35) as

(36) $$P[g_0(X, q_X) - g_0(X, p_X)] + P\left[\sum_{j=1}^{\infty} H_j g_j(X, q_j) - \sum_{j=1}^{\infty} H_j g_j(X, p_j)\right].$$

That the first expectation in (36) is nonnegative follows from the fact that g is strictly proper. From the assumption that P satisfies the law of total previsions in π,

$$P\left[\sum_{j=1}^{\infty} H_j g_j(X, q_j) - \sum_{j=1}^{\infty} H_j g_j(X, p_j)\right]$$

$$= P\left[P\left[\sum_{j=1}^{\infty} H_j g_j(X, q_j) - \sum_{j=1}^{\infty} H_j g_j(X, p_i)\middle|\pi\right]\right].$$

Using equation (5) and the same logic as in part (3.1), we obtain, for each i,

$$P\left[\sum_{j=1}^{\infty} H_j g_j(X, q_j) - \sum_{j=1}^{\infty} H_j g_j(X, p_j)\Big| H_i\right]$$
$$= P\left[H_i\{g_i(X, q_i) - g_i(X, p_i)\}|H_i\right]$$
$$= P\left[g_i(X, q_i)|H_i\right] - P\left[g_i(X, p_i)|H_i\right]$$
$$\geq 0,$$

where the final inequality follows because g_i is a proper scoring rule and $P(\cdot|H_i)$ is a finitely additive expectation for all i.

Therefore, since P satisfies the law of total previsions in π,

$$P\left[\sum_{j=1}^{\infty} H_j g_j(X, q_i) - \sum_{j=1}^{\infty} H_j g_j(X, p_i)\right] \geq 0,$$

which completes the proof of (36).

Corollary 3 is equivalent to the claim that for each set of rival forecasts, q_X and $\{q_j\}_{j=1}^{\infty}$, the second prevision in (36) is nonnegative, which was established in the proof of (3.2).

Acknowledgments. We thank Raphael Stern, Department of Statistics, Carnegie Mellon University, and two anonymous referees for helpful advice about this paper.

SUPPLEMENTARY MATERIAL

Infinite previsions and finitely additive expectations (DOI: 10.1214/14-AOS1203SUPP; .pdf). The expectation of a random variable X defined on Ω is usually defined as the integral of X over the set Ω with respect to the underlying probability measure defined on subsets of Ω. In the countably additive setting, such integrals can be defined (except for certain cases involving $\infty - \infty$) uniquely from a probability measure on Ω. Dunford and Schwartz [(1958), Chapter III] give a detailed analysis of integration with respect to finitely additive measures that attempts to replicate the uniqueness of integrals. Their analysis requires additional assumptions if one wishes to integrate unbounded random variables. We choose the alternative of defining integrals as special types of linear functionals. This is the approach used in the study of the Daniell integral. [See Royden (1963), Chapter 13.] Then the measure of a set becomes the integral of its indicator function. De Finetti's concept of prevision turns out to be a finitely additive generalization of the Daniell integral. (See Definition 10 in Appendix A.2.) We provide details on the finitely additive Daniell integral along with details about the meaning of infinite previsions and how to extend coherence$_1$ and coherence$_3$ to deal with ran-

dom variables having infinite previsions. Infinite previsions invariably arise when dealing with general sets of unbounded random variables.

REFERENCES

BERTI, P., REGAZZINI, E. and RIGO, P. (2001). Strong previsions of random elements. *Statistical Methods and Applications* **10** 11–28.

BERTI, P. and RIGO, P. (1992). Weak disintegrability as a form of preservation of coherence. *Journal of the Italian Statistical Society* **1** 161–181.

BERTI, P. and RIGO, P. (2000). Integral representation of linear functionals on spaces of unbounded functions. *Proc. Amer. Math. Soc.* **128** 3251–3258. MR1694449

BERTI, P. and RIGO, P. (2002). On coherent conditional probabilities and disintegrations. *Ann. Math. Artif. Intell.* **35** 71–82. MR1899945

CRISMA, L. and GIGANTE, P. (2001). A notion of coherent conditional prevision for arbitrary random quantities. *Stat. Methods Appl.* **10** 29–40.

CRISMA, L., GIGANTE, P. and MILLOSSOVICH, P. (1997). A notion of coherent prevision for arbitrary random quantities. *Journal of the Italian Statistical Society* **6** 233–243.

DE FINETTI, B. (1972). *Probability, Induction, and Statistics. The Art of Guessing*. Wiley, New York. MR0440638

DE FINETTI, B. (1974). *Theory of Probability*: *A Critical Introductory Treatment*, *Vol.* 1. Wiley, New York. MR0440640

DE FINETTI, B. (1975). *Theory of Probability*: *A Critical Introductory Treatment*, *Vol.* 2. Wiley, New York. MR0440641

DE FINETTI, B. (1981). The role of "Dutch Books" and of "proper scoring rules". *British J. Philos. Sci.* **32** 55–56. MR0618153

DUBINS, L. E. (1975). Finitely additive conditional probabilities, conglomerability and disintegrations. *Ann. Probab.* **3** 89–99. MR0358891

DUNFORD, N. and SCHWARTZ, J. (1958). *Linear Operators*. Wiley, New York.

GNEITING, T. (2011a). Making and evaluating point forecasts. *J. Amer. Statist. Assoc.* **106** 746–762. MR2847988

GNEITING, T. (2011b). Quantiles as optimal point forecasts. *International Journal of Forecasting* **27** 197–207.

HEATH, D. and SUDDERTH, W. (1978). On finitely additive priors, coherence, and extended admissibility. *Ann. Statist.* **6** 333–345. MR0464450

LEVI, I. (1980). *The Enterprise of Knowledge*. MIT Press, Cambridge, MA.

REGAZZINI, E. (1987). de Finetti's coherence and statistical inference. *Ann. Statist.* **15** 845–864. MR0888444

ROYDEN, H. L. (1963). *Real Analysis*. Macmillan, New York. MR0151555

SAVAGE, L. J. (1971). Elicitation of personal probabilities and expectations. *J. Amer. Statist. Assoc.* **66** 783–801. MR0331571

SCHERVISH, M. J., SEIDENFELD, T. and KADANE, J. B. (1984). The extent of nonconglomerability of finitely additive probabilities. *Z. Wahrsch. Verw. Gebiete* **66** 205–226. MR0749222

SCHERVISH, M., SEIDENFELD, T. and KADANE, J. (2008a). On the equivalence of conglomerability and disintegrability for unbounded random variables. Technical Report 864, Carnegie Mellon Univ., Pittsburgh, PA.

SCHERVISH, M. J., SEIDENFELD, T. and KADANE, J. B. (2008b). The fundamental theorems of prevision and asset pricing. *Internat. J. Approx. Reason.* **49** 148–158. MR2454836

SCHERVISH, M., SEIDENFELD, T. and KADANE, J. (2009). Proper scoring rules, dominated forecasts, and coherence. *Decis. Anal.* **6** 202–221.

SCHERVISH, M., SEIDENFELD, T. and KADANE, J. (2014). Supplement to "Dominating countably many forecasts." DOI:10.1214/14-AOS1203SUPP.

SEIDENFELD, T., SCHERVISH, M. J. and KADANE, J. B. (2009). Preference for equivalent random variables: A price for unbounded utilities. *J. Math. Econom.* **45** 329–340. MR2522457

M. J. SCHERVISH
J. B. KADANE
DEPARTMENT OF STATISTICS
CARNEGIE MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA 15213
USA
E-MAIL: mark@cmu.edu
        kadane@stat.cmu.edu

T. SEIDENFELD
DEPARTMENT OF STATISTICS
   AND DEPARTMENT OF PHILOSOPHY
CARNEGIE MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA 15213
USA
E-MAIL: teddy@stat.cmu.edu