Conglomerability, Disintegrability, and the Comparative Principle

Rush T. Stewart

Michael Nielsen *

February 8, 2021

1 Introduction

How stringent are standards of rational probability judgment? One prominent dividing line on this issue concerns additivity. Everyone agrees that numerically precise probabilities are required to be finitely additive. But many argue that they should meet the more restrictive standard of countable additivity as well. How are we to adjudicate the issue? Several approaches have been pursued. Our aim here is to present a result that connects some of these approaches, and, in doing so, to better understand the force of a recent argument for countable additivity due to Easwaran (2013). We have two main points. First, Easwaran's argument in favor of countable additivity should have little persuasive force on those permissive probabilists who have already made their peace with violations of conglomerability. As our result shows, Easwaran's main premise—the comparative principle—is strictly stronger than conglomerability. Second, with the connections between the comparative principle and other probabilistic concepts clearly in view, we point out that opponents of countable additivity can still make a case that countable additivity is an arbitrary stopping point between finite and full additivity. In the next section, we present our arguments in more detail, and in the final section we present the result on which our arguments are based.

2 Main Points

One important approach to justifying additivity requirements (along with other requirements) appeals to the concept of coherence, that is, the idea that one should always avoid accepting a set of bets that entails a sure loss. According to de Finetti, coherence is only a requirement for finite sets of bets. He claims that the extension of coherence to countable collections of bets is question begging (de Finetti, 1972, p. 91)—a controversial claim that has received some elaboration and defense (Howson, 2008). But, at least on the face of it, Dutch book arguments seem to support countable additivity as well. An issue with appealing to extensions of the coherence concept to justify the requirement of countable additivity is that the line

^{*}Thanks to Kenny Easwaran, Katie Steele, Borut Trpin, and three anonymous referees for helpful comments. Rush thanks the Center for Advanced Studies at LMU Munich for supporting this research.

of argument seems to generalize further in an analogous fashion to the justification of *full* additivity. We, like most others, take for granted that this is an unwelcome result. Among other things, full additivity rules out certain useful and natural distributions like the Lebesgue measure. So something is amiss with the coherence argument for countable additivity.

Another line of argument in favor of countable additivity focuses on the concept of conglomerability. Let (Ω, \mathcal{A}, P) be a finitely additive probability space, and define P(A|E) by the usual ratio formula when P(E) > 0. Let Π be a partition of Ω . We say that P is conglomerable in a positive partition¹ Π if

$$\forall A \in \mathcal{A} \inf_{E \in \Pi} P(A|E) \le P(A) \le \sup_{E \in \Pi} P(A|E).^2 \tag{1}$$

The probability P is conglomerable in all countable partitions if and only if P is countably additive. There are a few different motivations for conglomerability. First, failures of conglomerability allow for what has been called reasoning to a foregone conclusion (Kadane et al., 1996). Think of the cells of Π as the possible outcomes of an experiment. If P is not conglomerable in Π , then an agent with probabilities given by P could be sure that, no matter what the result of the experiment, the probability of some event A will be at least k after observing which cell $E \in \Pi$ is true, even though P(A) < k now. But such knowledge, some urge, should compel the agent to assign $P(A) \ge k$ before observing the outcome of the experiment. There is also a coherence argument for conglomerability making use of called-off bets. A bet on A conditional on E is called off (and the price of purchase refunded) if E is false, and proceeds as usual otherwise. If A witnesses that P is not conglomerable in Π , an agent is sure to lose if she makes a bet on A at price P(A) and a set of called-off bets on A conditional on $E \in \Pi$ at prices P(A|E) (Seidenfeld and Schervish, 1983, p. 410).

A similar issue arises for endorsing conglomerability for exactly all countable partitions as for extending coherence to exactly all countable sets of bets. Why should rationality require conglomerability for all *countable* partitions, but not all *uncountable* partitions? Some see it as a compelling requirement in general (Rescorla, 2015, p. 739). Whether we can have conglomerability for all uncountable partitions without mandating more than countable additivity depends on the theory of conditional probability that we adopt. Essentially, there is a trade off. With the theory of coherent conditional probabilities developed by de Finetti (1975) and Dubins (1975), countably additive probabilities will be non-conglomerable in some uncountable partition (Schervish et al., 2017). So, opponents of requiring countable additivity conclude, something is wrong with the conglomerability argument for countable additivity since the argument seems to generalize to support stronger forms of additivity which are universally acknowledged to be undesirable; conglomerability is mandatory only in the finite case. With the standard Kolmogorov theory of regular conditional probabilities, countably additive probabilities will be conglomerable in general. However, there are a number of anomalous features to cope with for Kolmogorov's theory including the Borel Paradox and the non-existence and impropriety of regular conditional probabilities (Seidenfeld, 2001; Schervish et al., 2017).³

 $^{^{1}}$ A partition is positive (with respect to P) if each cell of the partition has positive probability (according to P). Note that each cell of a partition can have positive probability only if the partition is countable.

 $^{^{2}}$ If Π is not positive, then some conditional probabilities will not be defined. We restrict our analysis to positive partitions, but note that it could be extended by appealing to primitive conditional probabilities.

³About the Kolmogorov approach of conditioning on sub-sigma-algebras rather than events, Kadane,

There is yet another, though similar, line of argument for countable additivity that begins with the notion of disintegrability. In a classic paper, Dubins investigated the connections between conglomerability and disintegrability, showing that the two are equivalent for expectations (linear previsions) defined on linear spaces (1975, Theorem 1). For the case of a probability function P defined on an algebra of events A, P is disintegrable in a positive partition Π if

$$\forall A \in \mathcal{A} \ P(A) = \sum_{E \in \Pi} P(A|E)P(E) \tag{2}$$

In other words, disintegrability implies the law of total probability. Since disintegrability is easily seen to imply conglomerability, P is disintegrable in all countable partitions only if P is countably additive. The motivations for conglomerability that we discussed above can also be used to motivate disintegrability. If (2) fails for some event A, then, by the Dubins result mentioned above, we have

$$\inf_{E \in \Pi} P(f|E) > P(f)$$

for some (bounded) random variable f, where P(f) and $P(f \mid E)$ are, respectively, the expectation and conditional expectation given E of f according to P. This is a failure of conglomerability for a random variable instead of an event. But the considerations that motivate conglomerability for events carry over to random variables. Consider reasoning to a foregone conclusion again, for instance. If one thinks that agents are required to have $P(f) \geq k$ whenever they are certain that their expectation for f will be at least k no matter what the outcome of an experiment is, then, by the Dubins result, one must require P to be disintegrable. The same goes for called-off bets. If P is not disintegrable, so that conglomerability fails for some f, then an agent is sure to lose if she makes a bet on f at price P(f) and a set of called-off bets on f conditional on $E \in \Pi$ at prices P(f|E).

Although disintegrability implies conglomerability, the two are not generally equivalent for probabilities on an algebra of events (Berti et al., 2017, Example 2). Along with Easwaran's comparative principle and conglomerability, disintegrability is a key concept in our formal result.⁴ Since disintegrability in a partition implies conglomerability in that partition, if P is not conglomerable in a partition, it is not disintegrable either. Just as with conglomerability, then, opponents of requiring countable additivity must accept some failures of disintegrability in countable partitions. Proponents of countable additivity, on the other hand, must accept failures of disintegrability (suitably generalized) in some uncountable partitions unless P satisfies full additivity—at least for coherent conditional probabilities (Dubins, 1975; Kadane et al., 1999; Howson, 2014). So, as with conglomerability, they confront the purported arbitrariness of stopping at requiring disintegrability for exactly countable partitions.

Easwaran advances two arguments in favor of countable additivity that cannot be extended to support full additivity (2013). The one that we focus on here is based on what Easwaran calls the comparative principle: if P is a probability function and Π is a measurable

Schervish, and Seidenfeld write, "This approach is unacceptable from the point of view of the statistician who, when given the information that A = B has occurred, must determine the conditional distribution of X_2 . A more reasonable approach is to consider the theory of conditional probability spaces as defined by (Dubins, 1975), among others" (1999, pp. 224–225).

⁴Among other things, the theorem below contributes to the literature a novel counterexample to the equivalence of disintegrability and conglomerability for probabilities over algebras of events.

partition, then, for every other probability function Q,

$$\neg \forall E \in \Pi \ P(E) < Q(E). \tag{3}$$

In words, no probability P can be dominated on a partition by another probability Q, in the sense that Q assigns every member of the partition greater probability than P. As Easwaran observes, "the only finitely additive, non-negative, normalized functions that are ruled out as probabilities by the Comparative Principle are the ones that violate countable additivity. This principle does nothing to rule out violations of full additivity $[\ldots]$ Countable additivity is not an arbitrary stopping point" (2013, p. 60). In other words, since the comparative principle motivates countable but not full additivity, this is an argument for exactly countable additivity. Easwaran states the comparative principle as an analysis of what probability is. For the subjective interpretation of probability, another way to put the point is that the comparative principle enjoins agents to avoid having credences that are dominated by another probability function in a partition in the sense of (3). Furthermore, since no (real-valued) probability assigns positive measure to each cell in an uncountable partition, every probability function avoids being dominated in such partitions. From this, it follows that the comparative principle is equivalent to requiring that probabilities be undominated in countable partitions.

While the comparative principle has been referenced a number of times in the literature (e.g., Elga, 2016; Arageorgis et al., 2017; Elliot, 2020; Hoek, MS), to the best of our knowledge, its relationships to standard probabilistic concepts other than countable additivity have received no attention outside of Easwaran's original discussion, in which the comparative principle is called "similar" and "dual" to but "distinct" from the concept of conglomerability (2013, p. 60, fn. 5). The result that we present in the next section relates the comparative principle to the classical probabilistic concepts of conglomerability and disintegrability in a precise way. We will show that the comparative principle is a strict strengthening of conglomerability and equivalent to disintegrability.

We want to make two points about the argument for countable additivity from the comparative principle. First, the theorem below facilitates the evaluation of the comparative principle as a norm on probabilistic judgments. Because the comparative principle is equivalent to disintegrability and implies conglomerability, reasons for and against mandating disintegrability and conglomerability are seen to have direct bearing on the status of the comparative principle. Since the comparative principle is strictly stronger than conglomerability for a single partition, and all proponents of relaxing countable additivity have reconciled themselves to failures of conglomerability, for such probabilists, the dialectical force of the argument for countable additivity on the basis of the comparative principle is marginal.⁵ And, similarly, since disintegrability implies but is not implied by conglomerability, and our result establishes that disintegrability is equivalent to the comparative principle, it should not be surprising if opponents of mandating countable additivity accept violations of the comparative principle with equanimity.

Second, when conceived of as a norm of personal probability, the comparative principle forbids being dominated in *any* (countable) partition. With this universal quantification,

⁵While the proof below proceeds differently, one can show that conglomerability in a partition follows from the comparative principle by taking $Q(\cdot) = P(\cdot|A)$ for any event A and using Bayes' theorem.

the comparative principle norm is equivalent to the norm requiring probabilities to be conglomerable in all countable partitions since both requirements are equivalent to countable additivity. In our opinion, evaluating the attractiveness of the comparative principle norm is not straightforward. Because of this, it is helpful to consider equivalent principles. Since the norm is equivalent to the conglomerability norm, considerations for and against conglomerability are manifestly relevant. Of particular interest are concerns about the arbitrary nature of requiring conglomerability for exactly countable partitions. While the comparative principle is supposed to present a case that countable additivity is not an arbitrary stopping point. that case is not so clear in light of the facts that avoiding dominance in all partitions is equivalent to avoiding dominance in all countable partitions and also equivalent to conglomerability in all countable partitions. Reflecting on the logical relationships the comparative principle bears to other probabilistic concepts helps us to better understand the set of relevant considerations.⁶ For example, since the problem of justifying stopping at conglomerability (and hence disintegrability) for countable partitions is especially salient, the trade offs involved in the theory of conditional probability are seen to be relevant considerations for the tenability of the comparative principle.

We do not anticipate that these points will dissuade those who find the comparative principle eminently compelling. However, even for such readers, the formal result below may still be of interest. But for those who, like Arageorgis et al. (2017, p. 12) and us, are more ambivalent about the merits of the comparative principle, we think the principle's connections to conglomerability should be illuminating.

3 Result

Theorem. Let (Ω, \mathcal{A}, P) be a probability space, and let Π be a positive measurable partition of Ω . (I) P obeys the comparative principle in Π (that is, (3) holds) if and only if P is disintegrable in Π (that is, (2) holds), and these imply that P is conglomerable in Π (that is, (1) holds). (II) The converse is not true: P may be both conglomerable and fail to obey the comparative principle.

Proof. For (I), start by assuming that P does not obey the comparative principle and that Q witnesses this, i.e. (3) is false. Then,

$$P(\Omega) = Q(\Omega) \geq \sum_{E \in \Pi} Q(\Omega \mid E) Q(E) = \sum_{E \in \Pi} Q(E) > \sum_{E \in \Pi} P(E) = \sum_{E \in \Pi} P(\Omega \mid E) P(E),$$

where the first inequality comes from the fact that finitely additive probabilities are countably superadditive. So P is not disintegrable in Π . Conversely, assume that P is not disintegrable in Π . Then, for some $A \in \mathcal{A}$,

$$P(A) > \sum_{E \in \Pi} P(A \mid E)P(E) = \sum_{E \in \Pi} P(A \cap E). \tag{4}$$

By (4) and countable superadditivity,

$$\sum_{E \in \Pi} P(E) = \sum_{E \in \Pi} \left(P(A \cap E) + P(A^c \cap E) \right) = \sum_{E \in \Pi} P(A \cap E) + \sum_{E \in \Pi} P(A^c \cap E) < P(A) + P(A^c) = 1.$$

⁶There are many more considerations than we discuss here (see, e.g., Savage, 1954; Dubins and Savage, 1965; de Finetti, 1972; Kelly, 1996; Seidenfeld, 2001; Nielsen and Stewart, 2019).

Enumerate the elements of Π by writing $\Pi = \{E_n : n \in \mathbb{N}\}$, and let $c = \sum_n P(E_n)$, so that c < 1. For all n, let $q_n = P(E_n) + 2^{-n}(1-c)$. Note that $q_n > P(E_n)$ and $\sum_n q_n = 1$. For each n, let $\omega_n \in E_n$, and define $Q(A) = \sum_n q_n \delta_{\omega_n}(A)$ for all $A \in \mathcal{A}$. Then Q dominates P on Π because $Q(E_n) = q_n$ for all n. So P does not obey the comparative principle. It's also worth noting that the dominating measure Q is countably additive. To see that let A_1, A_2, \ldots be a pairwise disjoint sequence in \mathcal{A} . Since each δ_{ω_n} is countably additive, we have

$$\sum_{n} q_n \delta_{\omega_n} \left(\bigcup_{m} A_m \right) = \sum_{n} \sum_{m} q_n \delta_{\omega_n} (A_m) = \sum_{m} \sum_{n} q_n \delta_{\omega_n} (A_m),$$

where the interchanging of the series is justified by the fact that $q_n \delta_{\omega_n}(A_m) \geq 0$ for all m and n.⁷

Finally, it's well known and not difficult to see that disintegrability implies conglomerability, so (I) is proved.

For (II), we modify a classic example due to Dubins. The idea is that there are two lotteries a and b over the natural numbers, exactly one of which will be activated. If a is activated, then n is chosen with probability 2^{-n} . If b is activated, then n is chosen with probability $\frac{2^{n-1}-1}{2^n(2^{n-1}+1)}$. The probability that a is activated is the same as the probability that b is activated. We note that b is a merely finitely additive lottery because $\sum_{n=1}^{\infty} \frac{2^{n-1}-1}{2^n(2^{n-1}+1)} < 1$, by comparison with $\sum_{n=1}^{\infty} 2^{-n} = 1$.

Example. Formally, let $\Omega = \{a, b\} \times \mathbb{N}$. Let $E_n = \{a, b\} \times \{n\}$ and $\Pi = \{E_n : n \in \mathbb{N}\}$. Let $A = \{a\} \times \mathbb{N}$ and $B = \{b\} \times \mathbb{N}$. Let P be a (merely) finitely additive probability defined on every subset of Ω satisfying

$$P\{(a,n)\} = \frac{1}{2^{n+1}}, \ P\{(b,n)\} = \frac{2^{n-1}-1}{2^{n+1}(2^{n-1}+1)}, \ P(A) = P(B).$$

Proposition 1. In the Example, P does not obey the comparative principle.

Proof of Proposition 1. Observe that $P(E_n) = P\{(a,n)\} + P\{(b,n)\} = \frac{1}{2^n+2} < \frac{1}{2^n}$, so P is dominated on Π by any probability Q on Ω such that $Q(E_n) = 2^{-n}$.

Proposition 2. In the Example, P is conglomerable in Π .

Proof of Proposition 2. To show that P is conglomerable in Π , it suffices to show that $\inf_n P(X \mid E_n) \leq P(X)$ holds for all $X \subseteq \Omega$, for then

$$1 - P(X) = P(X^c) \ge \inf_{n} P(X^c \mid E_n) = 1 - \sup_{n} P(X \mid E_n)$$

holds for all $X \subseteq \Omega$, leading to conglomerability. Let $X \subseteq \Omega$ be given.

The result is immediate if $X = \Omega$ or if $X \cap E_n = \emptyset$ for some n, so from now on assume that $X \neq \Omega$ and $X \cap E_n \neq \emptyset$ for all n.

Observe that

$$\frac{P\{(a,n)\}}{P(E_n)} = \frac{1}{2} + \frac{1}{2^n} \downarrow \frac{1}{2}$$
 (5)

⁷By definition, $\delta_{\omega_n}(A) = 1$ if $\omega_n \in A$ and 0 otherwise. Note that this construction of Q provides an elementary answer to a question about extending dominating measures raised by Easwaran (2013, fn. 7).

and

$$\frac{P\{(b,n)\}}{P(E_n)} = \frac{1}{2} - \frac{1}{2^n} \uparrow \frac{1}{2}.$$
 (6)

We now consider several cases.

Case 1: $A \subseteq X$. Because we are assuming that $X \neq \Omega$, we must have $X \cap E_n = \{(a, n)\}$ for some n, and we never have $X \cap E_n = \{(b, n)\}$ for any n. We now move to two sub-cases.

Case 1a: $X \cap E_n \neq E_n$ for infinitely many n. It follows that $X \cap E_n = \{(a, n)\}$ for infinitely many n. From (5) we get $\inf_n P(X \mid E_n) = 1/2 = P(A) \leq P(X)$.

Case 1b: $X \cap E_n = E_n$ for all but finitely many n. Let N be the largest n such that $X \cap E_n \neq E_n$. Then, $X \cap E_N = \{(a, N)\}$, so $P(X \mid E_N) = \frac{1}{2} + \frac{1}{2^N}$ by (5). It follows, also by (5), that $\inf_n P(X \mid E_n) = \frac{1}{2} + \frac{1}{2^N}$. By the fact that $X \supseteq \Omega \setminus \bigcup_{n=1}^N \{(b, n)\}$ and Lemma 1 below we have

$$P(X) \ge 1 - \sum_{n=1}^{N} P\{(b,n)\} \ge \frac{1}{2} + \frac{1}{2^N} = \inf_n P(X \mid E_n).$$

Case 2: $(a,n) \notin X$ for some n. The assumption that $X \cap E_n \neq \emptyset$ for all n implies that $X \cap E_n = \{(b,n)\}$ for some n. Let N be the least such n, so that, by (6), $\inf_n P(X \mid E_n) = 1/2 - 1/2^N$. Then, because $(a,n) \in X \cap E_n$ for all $n \leq N-1$, and using the facts that $\sum_{n=N}^{\infty} \frac{1}{2^{n+1}} = \frac{1}{2^N}$ and $\sum_{n=1}^{\infty} \frac{1}{2^{n+1}} = \frac{1}{2}$, we have

$$P(X) \ge \sum_{n=1}^{N-1} P(X \cap E_n) \ge \sum_{n=1}^{N-1} P\{(a,n)\} = \sum_{n=1}^{N-1} \frac{1}{2^{n+1}}$$
$$= \sum_{n=1}^{\infty} \frac{1}{2^{n+1}} - \sum_{n=N}^{\infty} \frac{1}{2^{n+1}} = \frac{1}{2} - \frac{1}{2^N} = \inf_n P(X \mid E_n).$$

This establishes that P is conglomerable in Π .

To complete the proof of the Theorem, it remains to state and prove the lemma used above.

Lemma 1. For all $N \in \mathbb{N}$,

$$1 - \sum_{n=1}^{N} \frac{2^{n-1} - 1}{2^{n+1}(2^{n-1} + 1)} \ge \frac{1}{2} + \frac{1}{2^{N}}.$$

Proof of Lemma 1. The result is immediate for N=1, so assume $N\geq 2$. Then,

$$\sum_{n=1}^{N} \frac{2^{n-1} - 1}{2^{n+1}(2^{n-1} + 1)} < \sum_{n=1}^{N} \frac{1}{2^{n+1}} - \frac{1}{4} < \frac{1}{4} \le \frac{1}{2} - \frac{1}{2^{N}},\tag{7}$$

where the first inequality follows from the inequality $\frac{2^{n-1}-1}{2^{n+1}(2^{n-1}+1)} < \frac{1}{2^{n+1}}$, which holds for all n, and, when n=1, is $0<\frac{1}{4}$. Multiplying by -1 and adding 1 to both sides of (7) yields the result.

References

- Arageorgis, A., J. Earman, and L. Ruetsche (2017). Additivity requirements in classical and quantum probability. *PhilSci-Archive eprints. URL http://philsci-archive. pitt. edu/13024*.
- Berti, P., E. Miranda, and P. Rigo (2017). Basic ideas underlying conglomerability and disintegrability. *International Journal of Approximate Reasoning* 88, 387–400.
- de Finetti, B. (1972). Probability, Induction, and Statistics. New York: John Wiley & Sons.
- de Finetti, B. (1990, originally published in 1975). Theory of Probability, Volume 1. John Wiley and Sons.
- Dubins, L. E. (1975). Finitely additive conditional probabilities, conglomerability and disintegrations. *The Annals of Probability* 3(1), 89–99.
- Dubins, L. E. and L. J. Savage (2014, originally published 1965). How to Gamble If You Must: Inequalities for Stochastic Processes. Mineola, New York: Dover Publications.
- Easwaran, K. (2013). Why countable additivity? Thought: A Journal of Philosophy 2(1), 53-61.
- Elga, A. (2016). Bayesian humility. Philosophy of Science 83(3), 305–323.
- Elliot, C. (2020). ET Jaynes's solution to the problem of countable additivity. *Erkenntnis Online First*, 1–22.
- Hoek, D. (MS). Chance and the continuum hypothesis. *Philosophy and Phenomenological Research*, Forthcoming.
- Howson, C. (2008). De Finetti, countable additivity, consistency and coherence. The British Journal for the Philosophy of Science 59(1), 1–23.
- Howson, C. (2014). Finite additivity, another lottery paradox and conditionalisation. Synthese 191(5), 989–1012.
- Kadane, J. B., M. J. Schervish, and T. Seidenfeld (1996). Reasoning to a foregone conclusion. *Journal of the American Statistical Association* 91 (435), 1228–1235.
- Kadane, J. B., M. J. Schervish, and T. Seidenfeld (1999). Statistical implications of finitely additive probability. In *Rethinking the Foundations of Statistics*, pp. 211–232. Cambridge University Press.
- Kelly, K. T. (1996). The Logic of Reliable Inquiry. Oxford University Press.
- Nielsen, M. and R. T. Stewart (2019). Obligation, permission, and bayesian orgulity. *Ergo, an Open Access Journal of Philosophy* 6(03), 58–70.
- Rescorla, M. (2015). Some epistemological ramifications of the Borel-Kolmogorov paradox. Synthese 192(3), 735–767.
- Savage, L. (1972, originally published in 1954). The Foundations of Statistics. New York: John Wiley and Sons.
- Schervish, M. J., T. Seidenfeld, and J. B. Kadane (2017). Non-conglomerability for countably additive measures that are not κ -additive. Review of Symbolic Logic 10(2), 284–300.
- Seidenfeld, T. (2001). Remarks on the theory of conditional probability: Some issues of finite versus countable additivity. In V. F. Hendricks (Ed.), *Probability Theory: Philosophy, Recent History and Relations to Science*. Synthese Library, Kluwer.

Seidenfeld, T. and M. J. Schervish (1983). A conflict between finite additivity and avoiding Dutch book. Philosophy of Science 50(3), 398–412.